

BIOMETRIE UND INFORMATIK

IN MEDIZIN UND BIOLOGIE

bisher
EDV in Medizin und Biologie

Gustav Fischer Verlag Stuttgart
Verlag Eugen Ulmer Stuttgart

Band 21
Heft 1/1990
ISSN 0934-9235

BIOMETRIE UND INFORMATIK

IN MEDIZIN UND BIOLOGIE

bisher EDV in Medizin und Biologie

Gustav Fischer Verlag Stuttgart
Verlag Eugen Ulmer Stuttgart

Schriftleitung

Prof. Dr. Hans Geidel, Stuttgart
Privatdozent Dr. Walter Lehmacher, München
Dipl.-Math. Joachim Vollmar, Mannheim

Herausgeber

P. Bauer (Köln) · H. Geidel (Stuttgart) · G. Grabner (Wien) · W. Lehmacher (München)
H. Rundfeldt (Hannover) · W. Stucky (Karlsruhe) · H. Thöni (Hohenheim) · J. Vollmar (Mannheim)

Wissenschaftlicher Beirat

R. Blomer (Frankfurt) · R. Haux (Heidelberg) · H. Haußmann (Hohenheim) · W. Köhler (Gießen)
M. Precht (Freising) · O. Richter (Braunschweig) · M. Schumacher (Freiburg) · E. Sonnemann (Trier)
H.-D. Unkelbach (Darmstadt) · H. F. Utz (Hohenheim)

Inhaltsverzeichnis 1/1990

Begrüßung zum Gedächtnis-Kolloquium für Univ.-Prof. Dr. med. et Dipl.-Math. Friedrich Wingert	U. Witting	2
Leben und Wirken von Friedrich Wingert	H. Götsche	4
Publikationen von Friedrich Wingert		7
Wachstumsprozesse in Biologie und Medizin	H.-J. Kretschmann	12
Biomathematik in Unterricht und Forschung	B. Schneider	22
Indexing Medical Information: The Role of SNOMED	D. J. Rothwell	31
SNOMED-Anwendungen	J. Berger/E. Hultsch/ R. Thurmayr	43

Redaktion:

Schriftleiter (verantwortlich im Sinne des Presserechts): Prof. Dr. H. Geidel,
c/o Verlag Eugen Ulmer, Postfach 70 05 61, 7000 Stuttgart 70, Tel. 47 28 80.

Verlag: Vertrieb und Werbung: Gustav Fischer Verlag GmbH & Co KG, 7000
Stuttgart 70, Wollgrasweg 49, Tel. (07 11) 45 50 38, Postgironkonto Stuttgart
13 556, Stuttgarter Bank 45 290. Herstellung und Anzeigen: Verlag Eugen
Ulmer GmbH & Co., Wollgrasweg 41, 7000 Stuttgart 70, Tel. (07 11) 45 07-0,
Telex 7-23 634, Postcheckkonto Stuttgart 74 63-700, Zürich 80-47072, Wien
1083.662. Deutsche Bank AG, Stuttgart, Kto. 14/76 878, Südwestbank AG,
Stuttgart, Kto. 21 000. Herstellung Siegfried Märkisch. Verantwortlich für den
Anzeigenteil: Dieter Boger, z. Z. ist die Anzeigenpreisliste Nr. 7 gültig. Anzei-
genschluß: Am 10. der Erscheinungsmonate Februar, Mai, August, November.

Druck: Kaiser-Druck GmbH, 7335 Salach, Schulstraße 11, Telefon (071 62)
70 11 + 70 12.

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt.
Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten.
Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verla-
ges in irgendeiner Form – durch Fotokopie, Mikrofilm oder andere Verfahren –
reproduziert oder in eine von Maschinen, insbesondere von Datenverarbei-
tungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der
Wiedergabe durch Vortrag, Funk- und Fernsehendung, im Magnettonverfah-
ren oder ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönli-
chen und sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen
oder Teilen daraus als Einzelkopien hergestellt werden.

Biometrie und Informatik in Medizin und Biologie ist eine deutsch/englischspra-
chige Zeitschrift, die theoretische und anwendungsbezogene Arbeiten aus dem
Gesamtgebiet der Biometrie und Informatik in Medizin und Biologie enthält. Die
Zeitschrift veröffentlicht sowohl Beiträge über neu entwickelte Ansätze als auch
über neue Anwendungsmöglichkeiten bekannter Methoden. Zusätzlich soll
dem tutoriellen Aspekt durch Übersichtsartikel Rechnung getragen werden.
Der Bezug zur Medizin oder Biologie darf dabei jedoch nicht vernachlässigt
werden. Jedes Heft enthält neben dem Leitartikel u. a.

- Übersichten zu bestimmten Themenkomplexen
- Originalien
- Informationen über Programme und Programmsysteme
- Aktuelle Informationen zu Veranstaltungen und Neuerscheinungen auf dem
Buchmarkt.

Die Zeitschrift wird zur Zeit referiert in den »Biological Abstracts« und »Current
Index to Statistics«.

Biometrie und Informatik in Medizin und Biologie erscheint vierteljährlich. Der
Bezugspreis beträgt jährlich DM 248,- (Einzelheft DM 75,-) zuzüglich Versand-
spesen. Vorzugspreis für Mitglieder der Deutschen Gesellschaft für Medizin-
ische Dokumentation und Statistik sowie der Deutschen Region der Internati-
onalen Biometrischen Gesellschaft DM 194,- zuzüglich Versandspesen. Bestel-
lungen nehmen jede Buchhandlung sowie die Verlage entgegen.

BIOMETRIE UND INFORMATIK

IN MEDIZIN UND BIOLOGIE

Bd. 21 (1990) Inhaltsverzeichnis

Editorial	1, 153
BAYERTZ, K.: Die Moralische Neutralität der Wissenschaft und ihre Grenzen	169
BERGER, J., HULTSCH, E. und THURMAYR, R.: SNOMED-Anwendungen	43
BRUNNER, E. und COMPAGNONE, D.: Die exakte Verteilung der Raviv-Statistik für kleine Stichprobenumfänge	113
DORDA, W.: Erfahrungen aus Entwicklung und Probetrieb eines „Statistischen Konsultationssystems“	57
FRITZ, H., UHLIG, K., SCHUNKERT, H., HAUX, R., WINTER, A., REPGES, R. und SIEBERTH, H.-G.: Analyse des Informatiksystems einer internistischen Klinik . . .	132
GÖTTSCHE, H.: Leben und Wirken von Friedrich Wingert	4
HOHLFELD, R.: Die Schöne Neue Welt der Humangenetik	181
KÖNIG, D. und KISON, H.-U.: EDV-gestütztes Verfahren der komplexen Ährenanalyse bei Getreide. II. Hard- und Software-Lösungen sowie Anwendererfahrungen	141
KOLLER, S.: Problemwandel in 6 Jahrzehnten biostatistischer Forschung – Persönliche Erinnerungen	232
KOZIOL, J. A.: On the span comparison matrix method for comparing protein or gene sequences	126
KRETSCHMANN, H.-J.: Wachstumsprozesse in Biologie und Medizin	12
LORENZ, R. J.: Die Arbeiten Siegfried Kollers zur Rassenhygiene in der Zeit von 1933 bis 1945	196
NOMAK, H., FERNER, U., UNKELBACH, H. D. und VOLLMAR, J.: Die besondere Verantwortung des Biometrikers in der Industrie	187
ROTHWELL, D. J.: Indexing Medical Information: The Role of SNOMED	31
SCHNEIDER, B.: Biomathematik in Unterricht und Forschung	22
WEINGART, P.: Politik und Vererbung	155
WELLEK, S.: Extended Tables of Exact Critical Values and Upper Tail Probabilities for Page's L-Statistic	71
WITTING, U.: Begrüßung zum Gedächtnis-Kolloquium für Univ.-Prof. Dr. med. et Dipl.-Math. Friedrich Wingert	2
WORTHHA, H.-P.: Varianzanalytische Anlagen mit zufälligen Besetzungszahlen – Assoziationsstrukturen zwischen qualitativen und quantitativen Merkmalen	94
Buchbesprechungen	106

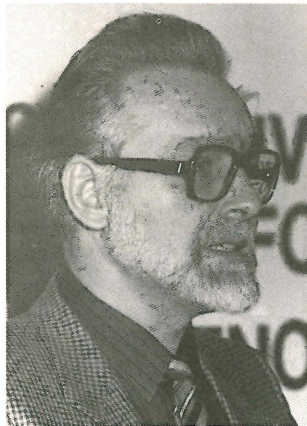
Editorial

Mit diesem Heft realisiert die Schriftleitung die Möglichkeit, in einem separaten Heft spezielle Themen aufzugreifen oder das Wirken einzelner Persönlichkeiten zu würdigen.

Ein Gedächtniskolloquium zu Ehren von Herrn Prof. Dr. Friedrich Wingert wurde am 10. 11. 1989 an der Universität Münster veranstaltet. In diesem Heft sind die schriftlichen Fassungen dieser Vorträge zusammengestellt. Nach einem Grußwort der Dekanin und einer Darstellung des Lebens und Wirkens von F. Wingert wurden im wissenschaftlichen Teil drei seiner Arbeitsschwerpunkte dargestellt: die Analyse von Wachstumsprozessen in Biologie und Medizin, die Arbeit beim Aufbau der Lehrangebote in Medizinischer Statistik und Informatik sowie die Arbeiten an dem System SNOMED (Systemized Nomenclature of Medicine). Die Vorträge des Kolloquiums werden um ein Schriftenverzeichnis ergänzt.

Begrüßung zum Gedächtnis-Kolloquium für Univ.- Prof. Dr. med. et Dipl.-Math. Friedrich Wingert

Prof. Dr. Ute Witting, Dekan der Medizinischen Fakultät



Sehr geehrte, liebe Frau Wingert,
sehr geehrte Freunde von Friedrich Wingert,
sehr geehrte Damen und Herren,

Friedrich Wingert hatte sich 1973 entschieden, den Ruf auf den Lehrstuhl für Medizinische Informatik und Biomathematik an der Medizinischen Fakultät der Westfälischen Wilhelms-Universität Münster anzunehmen und einen gleichzeitig an ihn ergangenen Ruf an die Universität Frankfurt abzulehnen. Er hat in Münster das Institut für Medizinische Informatik und Biomathematik als sein erster Direktor aufgebaut und bis zu seinem Tode am 29. Juli des vergangenen Jahres mit größtem persönlichem Engagement und Arbeitseinsatz geleitet.

Er wäre vor einigen Tagen 50 Jahre alt geworden. Aus diesem Anlaß veranstaltet die Medizinische Fakultät der

Westfälischen Wilhelms-Universität zusammen mit dem Institut für Medizinische Informatik und Biomathematik dieses Gedächtnis-Kolloquium.

Ich begrüße Sie herzlich im Namen der Medizinischen Fakultät der Westfälischen Wilhelms-Universität Münster und freue mich, daß so viele von Ihnen unserer Einladung gefolgt sind und an dieser Veranstaltung teilnehmen. Dies zeigt, wie sehr Sie Friedrich Wingert schätzten und wie gegenwärtig uns allen der Verstorbene ist. Die Referenten dieses Kolloquiums werden die Persönlichkeit und das wissenschaftliche Wirken von Herrn Wingert würdigen. Anzumerken bleibt die große Bereitwilligkeit, mit der die Vortragenden diese Aufgabe übernommen haben, und ich danke ihnen dafür, ganz besonders Doctor Rothwell, Clinical Professor of Pathology des Medical College of Wisconsin, der die weiteste Anreise hatte.

Danken möchte ich auch Herrn Dr. Dycka, Bayer AG, Herrn Vanderbeke, Hoechst AG, Herrn Dr. Plöger und Herrn Loos, Kassenzahnärztliche Vereinigung Westfalen-Lippe, Herrn Dr. Hartmann und Herrn Dr. Röpke, Schering AG, für ihr Interesse und ihre wertvolle Unterstützung dieser Veranstaltung, für deren arbeitsaufwendige Vorbereitung und Organisation besonders Herrn Priv.-Doz. Dr. Hultsch sehr herzlich zu danken ist.

Als Friedrich Wingert im Oktober 1973 die Leitung des Instituts für Medizinische Informatik und Biomathematik übernahm, wurden die im Studium der Medizin neuen Fächer „Biomathematik“ und „Medizinische Informatik“ eingeführt. In den ersten Jahren wurde die Bedeutsamkeit dieser Fächer für die Medizin unterschätzt. Wenn wir heute, etwa 15 Jahre später, zurückblicken, erkennen wir, welche Pionierarbeit Herr Wingert zu leisten hatte. Er

war dieser Aufgabe vollauf gewachsen, und es kommt uns zum Bewußtsein, was der Verstorbene für unsere Fakultät geleistet hat.

Eine Fakultät lebt zum großen Teil von der wissenschaftlichen Leistung der Hochschullehrer. Schon als Friedrich Wingert vor über 15 Jahren den Ruf nach Münster erhielt, war sein wissenschaftlicher Ruf gefestigt. Es ist zu seinen Lebzeiten in der Fakultät weitgehend unbekannt geblieben, mit welch hohem Einsatz, Erfolg und mit welch außergewöhnlicher internationaler Anerkennung seine wissenschaftlichen Tätigkeiten in Münster verknüpft sind.

Wie ich erfahren habe, beabsichtigen Sie, liebe Frau Wingert, eine Friedrich-Wingert-Stiftung zu gründen. Ich freue mich, daß so eine Institution gebildet wird, in der die Arbeit Ihres verstorbenen Mannes erfolgreich weitergeführt werden kann. Daß diese Arbeiten schon weitergeführt werden, können Sie unter anderem dem Veranstaltungsprogramm entnehmen: In der Nachmittagsveranstaltung, die dem gegenwärtigen Stand und den Perspektiven von SNOMED gewidmet ist, wird eine Demonstration des Programms durchgeführt.

Daß das Gedächtnis-Kolloquium für Friedrich Wingert in dieser Form stattfindet, zeigt, wie geachtet Herr Wingert während seines Wirkens war und in einem welch hohen Maße sein Andenken bewahrt wird.

Ich wünsche der Veranstaltung einen guten Verlauf und Ihnen allen bei intensiven Diskussionen einen erfolgreichen Erfahrungsaustausch.

Ich danke für Ihre Aufmerksamkeit.

Anschrift der Verfasserin: Prof. Dr. Ute Witting, Universität Münster, Institut für Arbeitsmedizin, Robert-Koch-Straße 51, D-4400 Münster, Telefon (02 51) 83-62 62

Leben und Wirken von Friedrich Wingert

H. Götsche

Friedrich Wingert wurde am 4. November 1939 in Darmstadt geboren. Er besuchte in Darmstadt auch die Volksschule und das Justus-Liebig-Gymnasium, an dem er 1959 das Abitur machte. Im Sommersemester des gleichen Jahres begann er in Frankfurt das Studium der Mathematik. Nach dem Vordiplom wechselte er das Studienfach und studierte ab dem Sommersemester 1963 Medizin.

Dieser so nüchtern vorgetragene Studienwechsel läßt den Schluß auf einen wesentlichen Zug in Friedrich Wingerts Wesen zu. Unzweifelhaft war er eine mathematische Begabung. Dieser folgte er, als er sich zum Mathematikstudium entschloß. Warum genügte ihm die Mathematik jedoch nicht? Nach meinem Eindruck und manchem persönlichen Gespräch beruhte dieser Entschluß auf zwei Wesenszügen in ihm: Sein umfassender Geist suchte den größeren Zusammenhang. Dieser schien ihm besonders gut herstellbar zu sein durch Anwendung der Methoden einer Wissenschaftsdisziplin auf eine andere. Warum aber nun gerade die Medizin? Vordergründig mag man zu dem Schluß kommen, daß gerade sie eine Fülle von Bezügen zu Naturwissenschaft und Mathematik hat und somit eine beträchtliche fachliche Nähe zur Mathematik besitzt. So plausibel eine solche Erklärung für die Entscheidung Friedrich Wingerts wäre, wir müssen sie aber doch mehr im privaten Bereich suchen: Die tiefe Verwurzelung in der Religion nämlich und die Begegnung mit seiner späteren Frau haben ihn hier in entscheidender Weise motiviert. Wer ihn gekannt hat, weiß, welche große Bedeutung diese Einflüsse für ihn hatten. Sie bieten uns im übrigen auch die Erklärung für die unbeirrbare Geradlinigkeit und Kompromißlosigkeit seines Wesens.

Das Doppelstudium brachte Friedrich Wingert dann auch bald intensiv mit Problemen aus dem Grenzgebiet von Mathematik und Medizin in Kontakt. Er entwickelte ein biometrisches Verfahren zur Oberflächenbestimmung von Hirnarealen, die einer direkten Beobachtung nicht zugänglich sind und durch Serienschnittverfahren nur mit großen methodischen Fehlern bestimmt werden konnten. Mit Hilfe des neuen Verfahrens konnten erstmals vergleichbare Messungen insbesondere bei innerhalb der Gehirnmasse liegenden Ansammlungen grauer Substanzen quantifiziert werden. Das Verfahren publizierte er 1968 in der Biometrischen Zeitschrift.

Das Medizin-Studium beendete Friedrich Wingert Anfang 1969 mit Promotion und der Note „summa cum laude“. Doktorvater war Hans-Joachim Kretschmann. Der Titel seiner Dissertation lautet: „Biometrische Analyse der Wachstumsfunktionen von Hirnteilen und Körpergewicht der Albinomäuse“. Die Dissertation wurde 1969 mit dem „Paul-Ehrlich-Preis für hervorragende Arbeiten des wissenschaftlichen Nachwuchses“ ausgezeichnet.

Nach der Promotion kehrte Friedrich Wingert nach Darmstadt zurück und war die folgenden 2 Jahre als wissenschaftlicher Mitarbeiter in der Abteilung für Numerik am Deutschen Rechenzentrum in Darmstadt tätig. Diese Zeit war wissenschaftlich äußerst fruchtbar. Friedrich Wingert setzte zugleich sein Mathematik-Studium an der Technischen Hochschule in Darmstadt fort und legte Ende 1970 die Diplomprüfung ab. Zudem gab ihm seine Tätigkeit am Deutschen Rechenzentrum die Gelegenheit, seine Zusammenarbeit mit Hans-Joachim

Kretschmann fortzusetzen. Friedrich Wingert entwickelte eine verallgemeinerte logistische Wachstumsfunktion, die sich als außerordentlich geeignet zur Darstellung der phylo- und ontogenetischen Entwicklung zentraler Strukturen erwies und die er 1971 in der Biometrischen Zeitschrift publizierte. Im gleichen Jahr veröffentlichte er zusammen mit Hans-Joachim Kretschmann eine zusammenfassende Monographie im Springer-Verlag.

Vom Deutschen Rechenzentrum in Darmstadt holte ihn 1971 Peter Leo Reichertz an das Department für Biometrie und Medizinische Informatik der Medizinischen Hochschule Hannover. Diese erteilte ihm am 7. 9. 1971 die *venia legendi* für die Fächer Medizinische Informatik und Biomathematik und ernannte ihn ab 1. 11. 1971 zum Oberassistenten.

Mit der Aufnahme seiner Tätigkeit in Hannover führte er seine bisherigen Forschungen zwar fort, aber sein Hauptinteresse wendete sich der Medizinischen Informatik zu. Er war maßgeblich in diesen Jahren am Aufbau des „Medizinischen Systems Hannover“ beteiligt. Im Januar 1971 wurde eine „Arbeitsgemeinschaft Medizinische Klartextdokumentation“ gegründet, der einige der größten deutschen Pathologischen Institute und zwei ausländische Institute angehörten. Ausgehend von einem früheren Versuch, Sektionsprotokolle systematisch zu analysieren, suchte er einen algorithmischen Zugang zur Erfassung und Beschreibung klartextlicher pathologisch-anatomischer Befunde.

Durch einen Forschungsaufenthalt in den National Institutes of Health in Bethesda im Jahre 1973 wurde er in seiner Überzeugung bestätigt, daß jede rechnerunterstützte Analyse von Klartexten nur unter Zugrundelegung einer mehrdimensionalen systematisch strukturierten Nomenklatur realisierbar ist. SNOP – die systematische Nomenklatur der Pathologie – erschien ihm ein geeigneter Ansatzpunkt. Zusammen mit P. Graepel stellte er in den USA eine deutsche Übersetzung von SNOP fertig.

Während dieses Forschungsaufenthaltes erhielt Friedrich Wingert gleichzeitig einen Ruf nach Frankfurt und einen Ruf nach Münster. Er entschied sich für Münster. Seit dem Jahr 1973 hat Friedrich Wingert als sein erster Direktor das Institut für Medizinische Informatik und Biomathematik der Westfälischen Wilhelms-Universität Münster bis zu seinem Tode am 29. Juli 1988 geleitet.

Der ausschlaggebende Aspekt bei Wingerts Entscheidung für Münster war, daß hier an der im Neubau befindlichen Klinik ein Medizinisches Rechenzentrum aufgebaut und durch das zu gründende Institut ein Medizinisches Informationssystem entwickelt werden sollte. Unter dieser Voraussetzung hat Friedrich Wingert das Institut für Medizinische Informatik und Biomathematik mit größtem persönlichen Engagement aufgebaut.

Der Name dieses Institutes enthält zwei Wissenschaftsbereiche. Friedrich Wingert hatte die *venia legendi* für beide Fächer. Sein Hauptinteresse galt jedoch der Medizinischen Informatik. So war es ihm eigentlich schon mehr eine Pflichtübung, zunächst die Biomathematik zu installieren: Dieses Fach war bei Gründung des Institutes ja gerade in die Medizinerausbildung aufgenommen worden. In den ersten Jahren wurde zusammen mit den Kollegen in Nordrhein-Westfalen ein für das gesamte Bundesgebiet grundlegendes Biomathematik-Buch und in kleinerem Kreis eine Fragensammlung erarbeitet und beim Springer-Verlag veröffentlicht, und es entstand ein Übungsbuch, das mit kleineren Änderungen auch heute noch benutzt wird. Friedrich Wingert war mehrere Jahre lang Sachverständiger beim Institut für medizinische und pharmazeutische Prüfungsfragen.

Wie schon gesagt, konnte Friedrich Wingert bei seiner Berufung nach Münster davon ausgehen, daß der Aufbau des Faches Medizinische Informatik mit der praktischen Realisierung eines Medizinischen Informationssystems verbunden sein würde. Dazu verfügte er ja aufgrund seiner Tätigkeit an der Medizinischen Hochschule Hannover über erhebliche Erfahrungen. Die in den Institutsstatuten niedergelegte Zuständigkeit für die EDV war von ihm nie anders interpretiert worden als Auftrag zu Errichtung und Betreuung eines Medizinischen Rechenzentrums. Diesen Plan verfolgte er mit großem Einsatz, unter anderem auch auf Landesebene. Die Vorstellungen der ihn umgebenden Institution bezüglich der aus der genannten Zuständigkeit folgenden Maßnahmen müssen aber wohl nicht allzu prä-

zise gewesen sein; denn Friedrich Wingert mußte erleben, daß sein Plan zu den Akten gelegt wurde und dem von ihm geführten Institut Stellen entzogen wurden. Friedrich Wingert, der kein Freund von Kompromissen war, hat diese nicht vorhersehbare Entwicklung so verarbeitet, daß er sich weitestgehend auf die Wissenschaft und dabei auf seine eigenen wissenschaftlichen Interessen konzentrierte.

Die SNOP, später SNOMED, die sich daran anschließenden linguistischen Probleme und Programmalgorithmen bestimmten bis zu seinem Tode seine wissenschaftliche Arbeit. Bald nach der Aufnahme seiner Tätigkeit in Münster erschienen die schon erwähnte Übersetzung von SNOP, wenig später Arbeiten zur medizinischen Linguistik, insbesondere zur Segmentierung der für die deutsche Sprache charakteristischen zusammengesetzten Ausdrücke, sowie eine Übersetzung von SNOMED, die von ihm gegenüber der amerikanischen Version auf den doppelten Umfang erweitert wurde. Parallel dazu liegt die Entwicklung von Software zur automatischen Indexierung nach der SNOMED und anschließenden Abbildung in die ICD, die er bis zur Serienreife vorantreiben konnte. In diesem groß angelegten Projekt nahm er Entwicklungen vorweg, die erst in jüngster Zeit als wissensbasierte Systeme Einzug in die allgemeine Informatik und die Computerlinguistik genommen haben.

Friedrich Wingert war Mitglied der Deutschen Gesellschaft für Medizinische Dokumentation, Informatik und Statistik (DGMDS) und der Gesellschaft für Informatik. Die DGMDS wählte ihn 1976 zum Vorsitzenden des Fachbereichs „Medizinische Informatik“, zum Leiter der Arbeitsgruppe „Klartextverarbeitung“ und zum Vorsitzenden der Zertifikatskommission „Medizinischer Informatiker“.

Einige Monate vor seinem Tod hatte er ein joint venture zusammen mit Doctor Rothwell (USA) und Doctor Côté (Canada) gegründet mit dem Ziel, die SNOMED-Lexika auch in Englisch und Französisch auf den deutschen Stand zu bringen und dann das Programm zur automatischen Indexierung zu vermarkten. Mein zuletzt mit ihm geführtes Gespräch, das auch sein letztes überhaupt sein sollte, handelte von dem Projekt der Übertragung seiner Programme auf den Mikrorechner.

Friedrich Wingert war zum Zeitpunkt seines Todes auf dem Höhepunkt seines Schaffens. Er war im besten Sinne besessen von seinen wissenschaftlichen Ideen. Die Beschäftigung damit war ihm Erfüllung. Er arbeitete am liebsten allein und mit einer Intensität, die für viele erst nach seinem frühen Tod offenbar wurde. Es war ihm aber vergönnt zu erfahren, wie seine Arbeit zunehmend nationale und internationale Anerkennung fand, wie er insbesondere weltweit als unumstrittene Erstautorität für die systematische medizinische Nomenklatur angesehen und geehrt wurde.

Viele Aufgaben lagen noch vor ihm – Aufgaben, die so leicht kein anderer übernehmen kann. Es ist trotzdem jetzt an uns zu versuchen, Friedrich Wingerts Werk fortzusetzen. Das Wirken Friedrich Wingerts wird vor allem all seinen Freunden, die hinter seiner mit größter Kompetenz und Konsequenz geleisteten fachlichen Arbeit auch den Menschen erkannten, dabei Vorbild und Antrieb zugleich sein.

Publikationen von Friedrich Wingert

- KRETSCHMANN HJ, WINGERT F: Über die Fehlergröße der Oberflächenbestimmung nach der Schnittserienmethode. *Z. wiss. Mikr.* 68 (1967), 93–114
- KRETSCHMANN HJ, WINGERT F, EMELE B: Quantitative Untersuchungen von Hirnzentren an ontogenetischen Reihen. *Zbl. ges. Neurol. Psychiat.* 192 (1968), 125
- WINGERT F, ERZ J: Eine neue numerische Methode zur Oberflächenbestimmung. *Biom. Zschr.* 10 (1968), 182–198
- KRETSCHMANN HJ, WINGERT F: Über die quantitative Entwicklung der Hippocampusformation der Albinomäuse. *J. Hirnforsch.* 10 (1968), 471–486
- WINGERT F: Biometrische Analyse der Wachstumsfunktionen von Hirnteilen und Körpergewicht der Albinomäuse. *J. Hirnforsch.* 11 (1969), 133–197
- KRETSCHMANN HJ, WINGERT F: Biometrische Analyse der Volumina des Corpus amygdaloideum an einer ontogenetischen Reihe von Albinomäusen. *Brain Res.* 12 (1969), 200–222
- KRETSCHMANN HJ, WINGERT F: Biometrische Analyse der Volumina des Striatum einer ontogenetischen Reihe von Albinomäusen. *Z. Anat. Entwickl.-Gesch.* 128 (1969), 85–108
- KRETSCHMANN HJ, WINGERT F: Biometrische Analyse der Volumina des Cortex piriformis einer ontogenetischen Reihe von Albinomäusen. *Z. Anat. Entwickl.-Gesch.* 129 (1969), 234–258
- WINGERT F: Eine Verallgemeinerung der logistischen Wachstumsfunktion. *Schriftenreihe des Deutschen Rechenzentrums*, Heft 9, Darmstadt 1969
- WINGERT F: Krankenhaus und Computer. *Bild der Wissenschaft* 5 (1970), 461–467
- WINGERT F: Sensitivitätsanalyse einiger bekannter statistischer Tests. *Schriftenreihe des Deutschen Rechenzentrums*, Heft 13, Darmstadt 1970
- KRETSCHMANN HJ, WINGERT F: Über die Größe und das Wachstum von Hirnzentren. *Verh. Anat. Ges.* 126 (1970), 181–190
- KRETSCHMANN HJ, SCHLEIFENBAUM L, WINGERT F: Quantitative Studies on the Postnatal Development of the Central Nervous System of *Cercopithecus aethiops*. *Proc. 3rd. Int. Congr. Primat.*, Zürich 1 (1970), 108–114
- WINGERT F: Eine Verallgemeinerung der logistischen Wachstumsfunktion. *Biometr. Zschr.* 13 (1971), 34–78
- KRETSCHMANN HJ, WINGERT F: Computeranwendungen bei Wachstumssproben in Biologie und Medizin. Berlin, Heidelberg, New York: Springer 1971
- WINGERT F: Computer-Prognosen. *Periskop* 2 (1972)
- WINGERT F: PAULA: Programm zur Auswertung logischer Ausdrücke. Plausibilitätskontrollen und Auswertung von Markierungsbelegen. *Meth. Inform. Med.* 11 (1972), 96–103
- WINGERT F: Klartextverarbeitung in der Pathologie. *Nds. Ärzteblatt* 45 (1972), 156–159
- ZILLES K, WINGERT F: Biometrische Analyse der Frischvolumina des Nucl. ruber einer ontogenetischen Reihe von Albinomäusen. *Z. Anat. Entwickl.-Gesch.* 138 (1972), 215–226
- ZILLES K, WINGERT F: Biometrische Analyse der Frischvolumina und der Anzahl der motorischen Neurone des Nucl. n. oculomotorii einer ontogenetischen Reihe von Albinomäusen. *Verh. Anat. Ges.* 67 (1973), 285–292
- ZILLES K, WINGERT F: Quantitative Studies of the Development of the Fresh Volumes and the Number of Neurones of the Nucl. n. oculomotorii of White Mice During Ontogenesis. *Brain Res.* 56 (1973), 63–75
- WINGERT F, RIES P: Pathologie-Befund-System. *Meth. Inform. Med.* 12 (1973), 150–155
- RÖTTGER P, WINGERT F, et al.: Konzeption und Organisation des AGK-Thesaurus. In: Siemens AG (Hrsg.): Symposium über Klartextanalyse in der Medizin (1). Erlangen: Siemens AG 1973, MC 12/1049, 52–60
- RIES P, WINGERT F, et al.: Die Erfassung von Biopsie-Berichten mit Hilfe des AGK-Thesaurus im Pathologischen Institut der Medizinischen Hochschule Hannover. In: Siemens AG (Hrsg.): Symposium über Klartextanalyse in der Medizin (1). Erlangen: Siemens AG 1973, MC 12/1049, 89–99

- KRETSCHMANN HJ, WINGERT F, ZILLES K: Biometrische Untersuchungen über die Hirnentwicklung bei *Tupaia belangeri*. Verh. Anat. Ges. 68 (1974), 389-397
- ZILLES K, KRETSCHMANN HJ, WINGERT F: Biometrische Analyse der Neuronenanzahl des Nucl. n. oculomotorii und des Nucl. n. trochlearis einer ontogenetischen Reihe von *Tupaia belangeri*. Verh. Anat. Ges. 68 (1974), 399-408
- WINGERT F: Textverarbeitung in der Medizin. EDV in Medizin und Biologie 5 (1974), 132-143
- WINGERT F: Word Segmentation and Morpheme Dictionary for Pathology Data Processing. In: Anderson J, Forsythe JM (Hrsg.): MEDINFO' 74. Uppsala: Almqvist u. Wiksell 1974, 915-921
- WINGERT F: Textverarbeitung in der Medizin. In: Siemens AG (Hrsg.): Symposium über Klartextanalyse in der Medizin (2). Erlangen: Siemens AG 1975, MC 12/1050, 17-43
- WINGERT F: Das Textverarbeitungssystem von PRATT. In: Reichertz PL, Holthoff G (Hrsg.): Methoden der Informatik in der Medizin. Berlin, Heidelberg, New York: Springer 1975, 216-223
- SCHLEICHER A, ZILLES K, KRETSCHMANN HJ, WINGERT F: Automatisierte Messungen verschiedener morphologisch erfaßbarer Parameter der Blutgefäße des Zentralnervensystems. Verh. Anat. Ges. 69 (1975), 155-161
- ZILLES K, SCHLEICHER A, KRETSCHMANN HJ, WINGERT F: Quantitative Untersuchungen über das Wachstum der Neuronen- und Gliazellanzahl in Kerngebieten verschiedener Species während der Ontogenese. Verh. Anat. Ges. 69 (1975), 505-515
- KRETSCHMANN HJ, SCHLEICHER A, WINGERT F, ZILLES K: Wachstumsparameter des Hirns und seiner Regionen bei Albinomäusen, *Tupaia belangeri* und SPF-Katze. Verh. Anat. Ges. 69 (1975), 517-525
- SCHLEICHER A, ZILLES K, WINGERT F, KRETSCHMANN HJ: Bestimmung der Anzahl der Zellen mit mehr als einem Nucleolus im histologischen Schnittpräparat. Microscopica acta 77 (1975), 316-330
- ZILLES K, SCHLEICHER A, WINGERT F: Spontaner Neuronenuntergang im Nucl. tract. mesencephali n. trigemini während der Perinatalzeit einer ontogenetischen Reihe von Albinomäusen. Gegenbaurs morph. Jb. 121 (1975), 289-299
- MARXKORS R, SOBANSKI U, WINGERT F: Erhebung zur befundbezogenen Planung von Zahnersatz. Dtsch. zahnärztl. Z. 30 (1975), 682-688
- WINGERT F: Kriterien für die Systemauswahl in der Medizin. Münch. med. Wschr. 117 (1975), 1751-1758
- WINGERT F, GRAEPEL P: Systematized Nomenclature of Pathology. Deutsche Übersetzung. Schriftenreihe des Instituts für Medizinische Informatik und Biomathematik der Universität Münster, Heft 1 (1975)
- HEINECKE A, HULTSCH E, REPGES R, WINGERT F: Examensfragen Biomathematik. Berlin, Heidelberg, New York: Springer 1975
- WINGERT F, FISCHER RJ: Medizinische Dokumentation. IBM (1975)
- KÜSEL W, RIES P, WINGERT F, RÖTTGER P, WESTERMANN, H: Ein variables Auswertungsprogramm für das Pathologie-Befund-System. In: Reichertz PL, Holthoff G (Hrsg.): Methoden der Informatik in der Medizin. Berlin, Heidelberg, New York: Springer 1975, 206-215
- WINGERT F: Einführung und Ausbau der elektronischen Datenverarbeitung an den medizinischen Einrichtungen der wissenschaftlichen Hochschulen des Landes Nordrhein-Westfalen. Ausschluß „ADV in der Medizin“ (Sondervotum), Münster 10. 7. 1975
- WINGERT F, GÖTTSCHE H: Einführung und Ausbau der elektronischen Datenverarbeitung an den medizinischen Einrichtungen der Universität Münster. Münster, 5. 11. 1975
- SCHLEICHER A, KRETSCHMANN HJ, WINGERT F, ZILLES K: Determination of the Number of Cells with Multiple Nucleoli in Histological Sections. National Bureau of Standards Special Publication 431. Proc. IV. Int. Congr. Stereol., Gaithersburg/USA 4.-9. 9. 1975, 371-374
- KOLLEGIUM BIOMATHEMATIK NW: Biomathematik für Mediziner. Berlin, Heidelberg, New York: Springer 1975, 2. Aufl.: Springer 1976
- HEINECKE A, HULTSCH E, NIENHAUS R, WINGERT F: Praktische Übungen Biomathematik. Schriftenreihe des Instituts für Medizinische Informatik und Biomathematik der Universität Münster, Heft 2 (1975), 2. Aufl. (1976)
- ZILLES K, SCHLEICHER A, KRETSCHMANN HJ, WINGERT F: Semiautomatic Morphometric Analysis of the Nuclear Development in the Nucl. n. oculomotorii of *Tupaia belangeri* During Ontogenesis. Anat. Embryol. 149 (1976), 15-28
- SCHLEICHER A, KRETSCHMANN HJ, WINGERT F, ZILLES K: Fehlermöglichkeiten und Fehlerkorrekturen beim Zählen von Zellen mit mehreren Nucleolen. Verh. Anat. Ges. 70 (1976), 917-922

- KRETSCHMANN HJ, SCHLEICHER A, WINGERT F, ZILLES K: Morphometrische Analyse von Wachstumsvorgängen. *Verh. Anat. Ges.* 70 (1976), 1035–1041
- ZILLES K, KRETSCHMANN HJ, NITSCHKE C, SCHLEICHER A, WINGERT F: Morphometrische Analyse der Entwicklung der Area striata bei *Tupaia belangeri*. *Verh. Anat. Ges.* 70 (1976), 329–337
- ZILLES K, SCHLEICHER A, WINGERT F: Quantitative Analyse des Wachstums der Frischvolumina limbischer Kerngebiete im Diencephalon und Mesencephalon einer ontogenetischen Reihe von Albinomäusen. I. Nucleus habenular. *J. Hirnforsch.* 17 (1976), 1–10
- ZILLES K, SCHLEICHER A, WINGERT F: Quantitative Analyse des Wachstums der Frischvolumina limbischer Kerngebiete im Diencephalon und Mesencephalon einer ontogenetischen Reihe von Albinomäusen. II. Corpus mamillare. *J. Hirnforsch.* 17 (1976), 11–20
- ZILLES K, SCHLEICHER A, WINGERT F: Quantitative Analyse des Wachstums der Frischvolumina limbischer Kerngebiete im Diencephalon und Mesencephalon einer ontogenetischen Reihe von Albinomäusen. III Nucleus interpeduncularis. *J. Hirnforsch.* 17 (1976), 21–29
- WINGERT F, GEISLER U: Anleitung zum Erstellen von Programmen für Daten in Form binärer Bäume. Schriftenreihe des Instituts für Medizinische Informatik und Biomathematik der Universität Münster, Heft 3 (1976)
- WINGERT F: Medical Language Data Processing. In: Reichertz PL, Goos G (eds.): *Informatics and Medicine*. Berlin, Heidelberg, New York: Springer 1977, 579–646
- WINGERT F: Medizinische Informatik. In: Zenker R et al. (Hrsg.): *Chirurgie der Gegenwart*, Band I. München, Wien, Baltimore: Urban & Schwarzenberg 1977
- KRETSCHMANN HJ, SCHLEICHER A, WINGERT F, ZILLES K: Quantitativ-mikroskopische Morphologie motorischer Hirnregionen während der Ontogenese von *Tupaia belangeri* und SPF-Katze. *Verh. Anat. Ges.* 71 (1977), 133–138
- ZILLES K, KRETSCHMANN HJ, WINGERT F: Ergebnisse quantitativ-mikroskopischer Untersuchungen der Morphologie des visuellen Systems während der Ontogenese von *Tupaia belangeri*. *Verh. Anat. Ges.* 71 (1977), 139–146
- WINGERT F: Medizinische Informatik. *Med. Klin.* 72 (1977), 1281–1287
- WINGERT F: Morphosyntaktische Zerlegung von Komposita der medizinischen Sprache. *Meth. Inform. Med.* 16 (1977), 248–255
- GÖTTSCHE H, WINGERT F: Basisdokumentation für die medizinischen Einrichtungen der Universität Münster. Schriftenreihe des Instituts für Medizinische Informatik und Biomathematik der Universität Münster, Heft 4 (1977)
- WINGERT F: A Model for Segmentation of Medical Compound Word Forms. In: Shires DB, Wolf H (eds.): *MEDINFO' 77*. Amsterdam, New York, Oxford: North-Holland 1977, 241–245
- FORCK G, WINGERT F: Kontaktterkrankungen des Zahnarztes und seiner Helfer. I. Teil: *ZWR* 86 (1977), 352–360. II. Teil: *ZWR* 86 (1977), 412–416
- WINGERT F: Morphosyntactical Analysis of Medical Compound Word Forms. In: Schneider W, Sägval Hein AL, (eds.): *Computational Linguistics in Medicine*. Amsterdam, New York, Oxford: North-Holland 1977, 79–89
- FISCHER RJ, WINGERT F: Ein allgemein verwendbares PI/1-Makro für Message-Processing und Fehlerbehandlung in Programmen unter dem Information Management System (IMS). In: Reichertz PL, Schwarz B (Hrsg.): *Informationssysteme in der medizinischen Versorgung*. Stuttgart, New York: Schattauer 1978, 291–297
- WINGERT F (Hrsg.): *Klartextverarbeitung*. Berlin, Heidelberg, New York: Springer 1978
- WINGERT F: Klartextverarbeitung in der Medizin. In: Wingert F (Hrsg.): *Klartextverarbeitung*. Berlin, Heidelberg, New York: Springer 1978, 1–20
- JACOB W, SCHEIDA D, WINGERT F (Hrsg.): *Tumor-Histologie-Schlüssel ICD-O-DA*. International Classification of Diseases for Oncology. Berlin, Heidelberg, New York: Springer 1978
- KRETSCHMANN HJ, SCHLEICHER A, WINGERT F, ZILLES K, LÖBLICH HJ: Human Brain Growth in the 19th and 20th Century. *J. Neurol. Sciences* 40 (1979), 169–188
- WINGERT F: *Medizinische Informatik*. Stuttgart: Teubner 1979
- WINGERT F: Contributions to Language Data Processing. *J. Clin. Comp.* 8 (1979), 155–180
- WINGERT F: Medical Linguistics: A Review. In: Lindberg DAB, Kaihara S (eds.): *MEDINFO' 80*. Amsterdam, New York, Oxford: North-Holland 1980, 1321–1331
- WINGERT F: *ASPECT-Manual*. Berlin: Schering AG 1980

- BÜCHNER T, URBANITZ D, WINGERT F: Chemotherapie und Immuntherapie zur Erhaltung der kompletten Remission bei akuter myeloischer Leukämie – eine monozentrische kontrollierte klinische Studie in der Planung. *Verh. Dtsch. Ges. Innere Med.* 86 (1980), 1296–1303
- WINGERT F: *Medical Informatics*. Berlin, Heidelberg, New York: Springer 1981
- WINGERT F: *ASPECT-Manual*. Englische Ausgabe. Berlin: Schering AG 1981
- WINGERT F: Methoden der Medizinischen Linguistik. In: Adlassnig KP, Dorda W, Grabner G (Hrsg.): *Medizinische Informatik*. Wien, München: Oldenbourg 1981, 15–47
- HEINECKE A, HULTSCH E, WINGERT F: *Praktische Übungen Biomathematik*. Schriftenreihe des Instituts für Medizinische Informatik und Biomathematik der Universität Münster, Heft 2, 5. Aufl. (1981)
- SAUER B, KAMMRADT G, KRAUTHAUSEN I, KRETSCHMANN HJ, LANGE HW, WINGERT F: Qualitative and Quantitative Development of the Visual Cortex in Man. *J. Comp. Neurol.* 214 (1983), 441–450
- WINGERT F: Allgemeine Einführung in die EDV mit Begriffserläuterungen. In: KZV-WL (Hrsg.): *Neue Organisationsformen in der Zahnarztpraxis* 1984, 9–14
- WINGERT F: *SNOMED*. Systematisierte Nomenklatur der Medizin. Numerischer Index. Berlin, Heidelberg, New York, Tokyo: Springer 1984
- WINGERT F: *SNOMED*. Systematisierte Nomenklatur der Medizin. Alphabetischer Index. Berlin, Heidelberg, New York, Tokyo: Springer 1984
- WINGERT F: *SNOMED*. Manual. Berlin, Heidelberg, New York, Tokyo: Springer 1984
- HULTSCH E, WINGERT F: Modelle zur Beschreibung des Wachstumsverhaltens von Organen und Organteilen bei Versuchstieren. *Z. Versuchstierkd.* 26 (1984), 250–251
- WINGERT F: Automated Indexing Based on SNOMED. *Meth. Inform. Med.* 24 (1985), 27–34
- WINGERT F: Reduction of Redundancy in a Categorized Nomenclature. In: Côté RA, Protti DJ, Scherrer JR (eds.): *Role of Informatics in Health Data Coding and Classification Systems*. Amsterdam, New York, Oxford: North-Holl. 1985, 191–202
- WINGERT F: Morphologic Analysis of Compound Words. *Meth. Inform. Med.* 24 (1985), 155–162
- WINGERT F: Contributions to Automated Indexing Based on SNOMED. In: Roger FH, Grönroos P, Tervo-Pellika R, O'Moore R (eds.): *Medical Informatics 85*. Berlin, Heidelberg, New York: Springer 1985, 211–215
- WINGERT F: *ASPECT-Manual*. 2. Auflage. Berlin: Schering AG 1985
- HEINECKE A, HULTSCH E, WINGERT F: *Praktische Übungen Biomathematik*. Schriftenreihe des Instituts für Medizinische Informatik und Biomathematik der Universität Münster, Heft 2, 7. Aufl. (1985)
- WINGERT F: An Indexing System for SNOMED. *Meth. Inform. Med.* 25 (1986), 22–30
- KRETSCHMANN HJ, KAMMRADT G, KRAUTHAUSEN I, SAUER B, WINGERT F: Brain Growth in Man. *Bibl. Anatomica* 28 (1986) 1–26
- KRETSCHMANN HJ, KAMMRADT G, KRAUTHAUSEN I, SAUER B, WINGERT F: Growth of the Hippocampal Formation in Man. *Bibl. Anatomica* 28 (1986), 27–52
- WINGERT F: Automatische Indexierung in die ICD. In: Ehlers CT, Beland H (Hrsg.): *Perspektiven der Informationsverarbeitung in der Medizin*. Berlin, Heidelberg, New York, London, Paris, Tokyo: Springer 1986, 215–218
- WINGERT F, FISCHER RJ, OSADA N: Untersuchungen zur Verminderung des Aufwands bei automatischer Diagnoseverschlüsselung. In: Ehlers CT, Beland H (Hrsg.): *Perspektiven der Informationsverarbeitung in der Medizin*. Berlin, Heidelberg, New York, Paris, Tokyo: Springer 1986, 232–235
- WINGERT F: Automated Mapping of ICD into SNOMED. In: Orthner HF (ed.): *The 10th Annual Symposium on Computer Applications in Medical Care*. Washington DC: IEEE Computer Soc. Press 1986, 81–84
- WINGERT F: Introduction to Medical Linguistics and SNOMED. *Unterlagen zum SCAMC-Tutorial*, Washington 1986
- GÖHRING R, HAAS P, HOFFMANN WD, KLAR R, WILDE E, WINGERT F: Empfehlung zur Dokumentation und Auswertung von Diagnosen in Krankenhäusern. *Schriftenreihe der GMDs*, Nr. 8
- WINGERT F: Automated Indexing Based on SNOMED. In: Salamon R, Blum B, Jørgensen M (eds.): *MEDINFO 1986*. Amsterdam, New York, Oxford, Tokyo: North-Holland 1986, 1149
- WINGERT F: Automated Indexing of SNOMED Statements into ICD. *Meth. Inform. Med.* 26 (1987), 93–98
- WINGERT F: Medical Linguistics: Automated Indexing into SNOMED. *Critical Reviews in Medical Informatics* 1 (1988), 333–403

- SPRANGER H, WINGERT F, PLÖGER CTH: Erhebungen über befundbezogene Ausführung von Zahnersatzformen Krankenversicherter in Westfalen-Lippe 1987. ZM 78 (1988), 1976-1989
- WINGERT F: Grundlagen der Indexierung medizinischer Diagnosen und Therapien. In: Wille R (Hrsg.): Klassifikation und Ordnung. Studien zur Klassifikation, Bd. 19. Frankfurt: Indeks-Verlag 1989, 165-178
- WINGERT F, ROTHWELL D, CÔTÉ R: Automated Indexing into SNOMED and ICD. In: Scherrer JR, Côté RA, Mandil SH (eds.): Computerized Natural Medical Language Processing for Knowledge Representation. Amsterdam, New York, Oxford, Tokyo: North-Holland 1989, 201-239

Wachstumsprozesse in Biologie und Medizin

H.-J. Kretschmann

Zusammenfassung

FRIEDRICH WINGERT entwickelte mathematische Grundlagen für die computerunterstützten Anwendungen bei Wachstumsvorgängen in Biologie und Medizin. Er verallgemeinerte die 3-parametrische logistische Wachstumsfunktion (VERHULST 1845, 1847) zu einer 4- und 5-parametrischen logistischen Wachstumsfunktion. Die Programme Logi und Komb (WINGERT 1969, 1971) wurden besonders im Bereich neuroontogenetischer Prozesse eingesetzt und haben neue Einsichten für die Hirnforschung erbracht.

Schlüsselwörter: generalisierte logistische Wachstumsfunktion; Hirnwachstum.

Summary

FRIEDRICH WINGERT developed the mathematical principles for the computerassisted analysis of growth processes in biology and medicine. He generalized the 3-parametric logistic growth function (VERHULST, 1845, 1847) to a 4- and 5-parametric one. The programs Logi and Komb (WINGERT 1969, 1971) were applied mainly for neuroontogenetic processes and have given new insight for brain research.

Key words: generalization of the logistic growth function; brain growth.

Die praktische Bedeutung der Wachstumsprozesse in Biologie und Medizin reicht von der Landwirtschaft, der Tiermedizin bis hin zur Kinderheilkunde. Das Wachstum ist für alle Lebewesen existenziell. Es gehört wie der Stoffwechsel, die Atmung, die Vermehrung und die Kommunikation mit der Umwelt zu den grundlegenden Eigenschaften der Lebewesen. Vor 190 Jahren hat der Engländer MALTHUS seine Zeitgenossen mit der These erschreckt, daß die natürliche Vermehrung der Bevölkerung nach einer geometrischen Progression und der Ertragszuwachs des Bodens nach einer arithmetischen Progression erfolgen würde. Die Kurvenverläufe des biologischen Wachstums sind in der Regel jedoch komplizierter, als daß sie mit einer geometrischen oder arithmetischen Reihe beschrieben werden könnten.

Die ersten Ansätze für eine mathematische Analyse von S-förmigen Wachstumsverläufen stammen von dem Belgier VERHULST, der 1845 das Modell einer 3-parametrischen Wachstumsfunktion entwickelte:

$$(1) \quad y = \frac{P_1}{1 + \exp(P_2 + P_3 \cdot t)}$$

Hierbei sind P_1 , P_2 und P_3 die Parameter, die Zeit t ist die unabhängige, die Wachstumsgröße y die (im statistischen Sinne) abhängige Variable. Der Rechenaufwand für die Bestimmung der Parameter der logistischen Wachstumsfunktion war vor Einführung der elektronischen Datenverarbeitung erheblich. Eine gewisse Vereinfachung brachte die Einführung der Logits durch BERKSON (1944). In der Erforschung der Wachstumsprozesse wurden kurz nach dem zweiten Weltkrieg in der Biologie und Medizin folgende Ansätze verfolgt:

1. Viele Autoren begnügten sich mit Freihandkurven, um den Verlauf des Wachstums anzugeben. Dieses Verfahren gibt jedoch nur grobe Informationen über das Wachstum und schließt eine wissenschaftlich-statistische Analyse aus.
2. In die S-förmigen Datenverläufe wurden lineare Funktionen hineingelegt, die jedoch zu falschen Schlüssen führen mußten, weil sie keine Aussagen über den gesamten Datenverlauf ermöglichen.
3. Es wurde die Abhängigkeit der Organgewichte von den Körpergewichten oder von ähnlichen Gesamtgrößen untersucht und diese Beziehung „allometrisches Wachstum“ genannt. Das Wachstum ist jedoch immer eine zeitabhängige Größe. Deshalb sind die Beziehungen Organgewicht zu Körpergewicht besser als Korrelationen einzuordnen.

1966 begann FRIEDRICH WINGERT seine wissenschaftliche Laufbahn im neuroanatomischen Bereich. Der Autor war der Betreuer der Dissertation von FRIEDRICH WINGERT und in diesem Sinne auch Augenzeuge seiner wissenschaftlichen Aktivität. FRIEDRICH WINGERT analysierte tiefgründig und besaß eine unbeschreibliche Energie, schwierige biologisch-medizinische Probleme zu lösen. Seine, mit Summa cum laude ausgezeichnete Promotion erbrachte die computergestützte Methode, S-förmige Wachstumsprozesse mit der logistischen Wachstumsfunktion zu analysieren. FRIEDRICH WINGERTS anschließende Tätigkeit am Deutschen Rechenzentrum in Darmstadt führte zur Verallgemeinerung der logistischen Wachstumsfunktion, mit der auch komplizierte Wachstumsverläufe gut analysiert werden konnten. Bei der Verallgemeinerung der logistischen Wachstumsfunktion fügte FRIEDRICH WINGERT einen quadratischen und einen kubischen Exponenten ein:

$$(2) \quad y = \frac{P_1}{1 + \exp(P_2 + P_3 \cdot t + P_4 \cdot t^2 + P_5 \cdot t^3)}$$

Bereits PEARL und REED (1923) erzeugten schiefe Wachstumskurven durch Einfügen einer größeren Anzahl von Parametern in den Exponenten. Der Rechenaufwand der Anpassung der zusätzlichen Parameter mit Hilfe der Gaußschen Methode der kleinsten Quadrate war damals jedoch so aufwendig, daß die Berechnung erst mit der Einführung leistungsfähiger Rechner und der Entwicklung effizienter Programme ermöglicht wurden.

Das Programm Logi (WINGERT 1969 b, 1971) verwendet ebenfalls die Methode der kleinsten Quadrate. Nach einer Anfangsschätzung der Parameter P_1 , P_2 , P_3 wird die beste Kurve für diese 3 Parameter gesucht. Dann werden P_1 , P_2 , P_3 festgehalten und dazu der beste Parameter P_4 gesucht. Schrittweise wird anschließend der beste Parameter P_5 iteriert. Der Parameter P_1 gibt den im „Endzustand“ erreichten Idealwert der Wachstumskurve an. Eine biologisch wichtige Größe des Wachstums ist die Halbwertszeit. Darunter wird die Zeit verstanden, nach der die Wachstumsgröße y den halben Idealwert P_1 erreicht.

Außer diesen Werten gibt das Programm Logi auch die Reifegrade der logistischen Wachstumsfunktion an. Die Funktionen der Reifegrade erhält man, wenn die Wachstumsfunktionen (1) und (2) durch P_1 dividiert werden. Die Funktion der Reifegrade strebt gegen 1 und enthält die gesamte Information der Wachstumsform – mit Ausnahme des Parameters P_1 . Der Vergleich der Reifegrade erlaubt deshalb bessere Einblicke in die Wachstumsdynamik als der Vergleich der absoluten Wachstumswerte (Abb. 1–8).

Die Wachstumsrate der Reifegrade ist die Ableitung der Funktion der Reifegrade. Die Wachstumsrate der Reifegrade beschreibt die Änderungstendenz der Funktion der Reifegrade an jeder Stelle (Abb. 6). Das Programm Komb berechnet die Parameter von zwei 3-parametrischen Wachstumsfunktionen, die sich überlagern. Die erste logistische Wachstumsfunktion (in Abb. 9 gepunktet und mit (1) beschriftet) vergrößert sich monoton, während die zweite logistische Wachstumsfunktion (in Abb. 9 gepunktet und mit (2) beschriftet) monoton abnimmt.

Mehr als hundert Hirnstrukturen von Tier und Mensch wurden mit FRIEDRICH WINGERTS Programm Logi untersucht (KRETSCHMANN, WINGERT 1971, KRETSCHMANN et al. 1979, 1986 a,

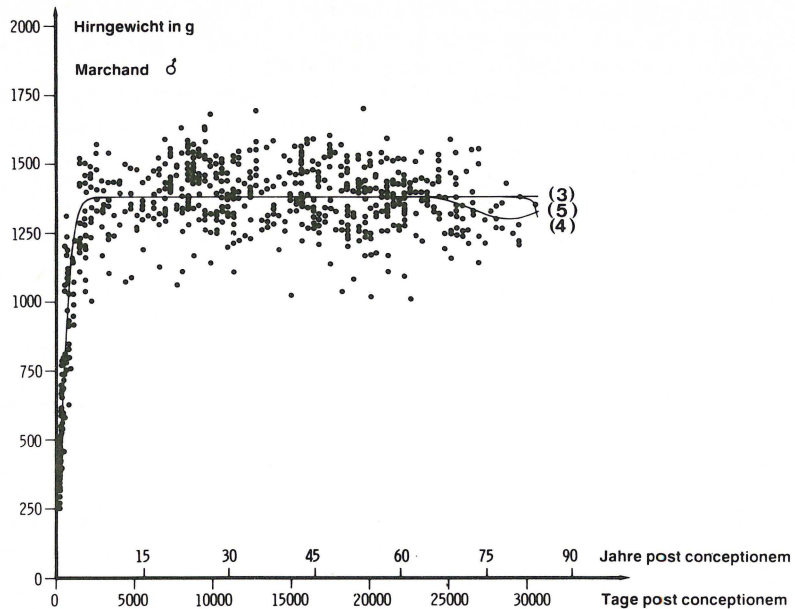


Abbildung 1 a 716 männliche Hirngewichte vom Menschen wurden gegen ihr postkonzeptionelles Alter aufgetragen und mit der 3-, 4- und 5-parametrischen logistischen Wachstumsfunktion approximiert. Die Daten stammen von MARCHAND (1902).

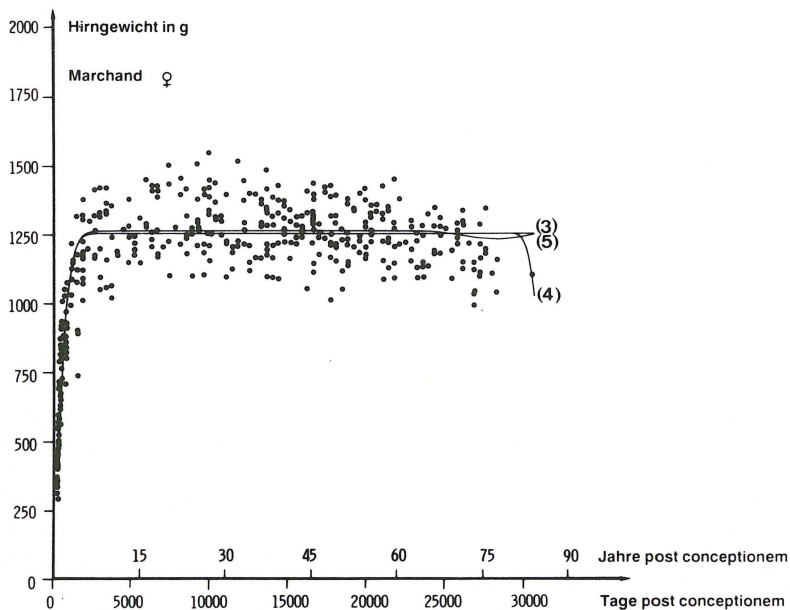


Abbildung 1 b Die logistischen Wachstumsfunktionen wurden 452 weiblichen Hirngewichten und ihrem Lebensalter angepaßt. Die Daten wurden von Marchand in den Jahren um die Jahrhundertwende bestimmt.

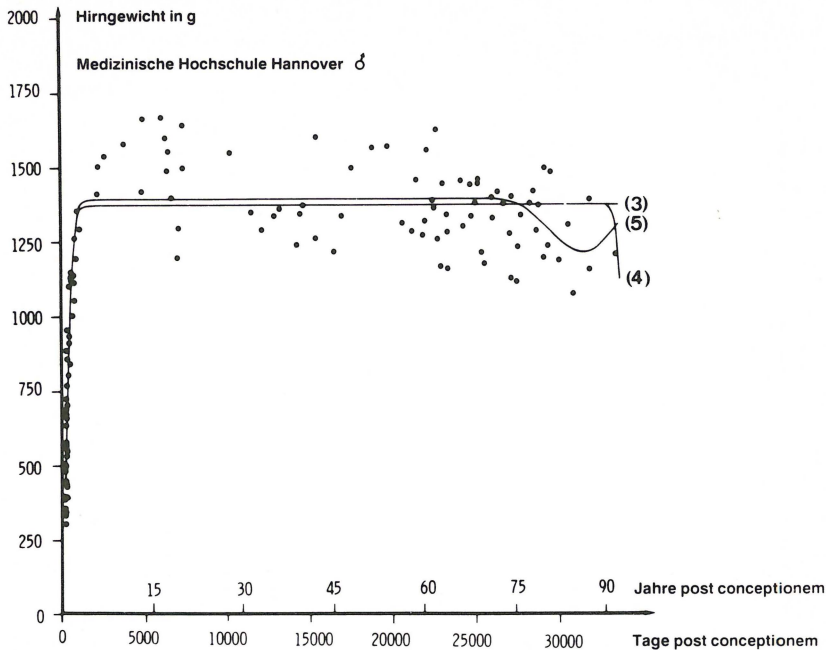


Abbildung 2 a Die 3-, 4- und 5-parametrischen logistischen Wachstumsfunktionen wurden 129 männlichen Hirngewichten vom Menschen und ihrem Lebensalter angepaßt. Die Daten wurden in den Jahren 1974 – 1978 an der Medizinischen Hochschule Hannover gewonnen.

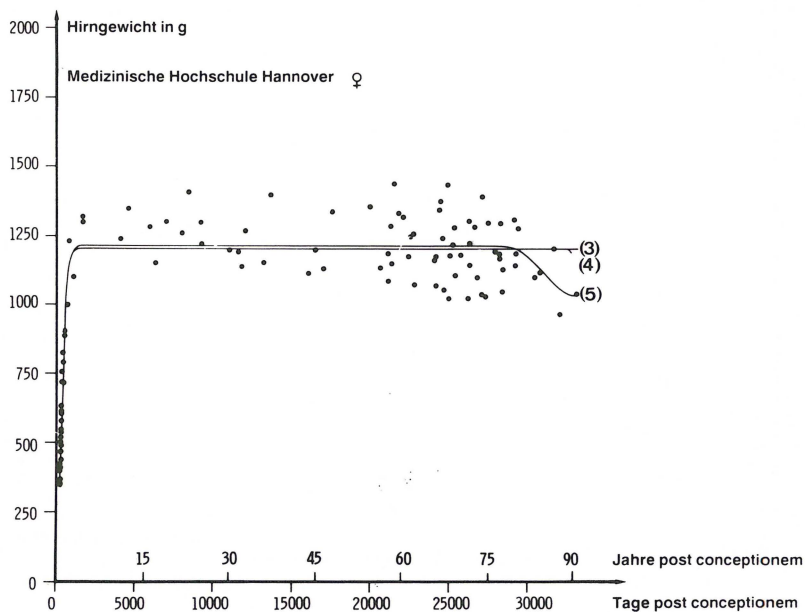


Abbildung 2 b Die logistischen Wachstumsfunktionen einer Stichprobe von 104 weiblichen Hirngewichten und ihrem Lebensalter aus denselben Jahren an der Medizinischen Hochschule Hannover.

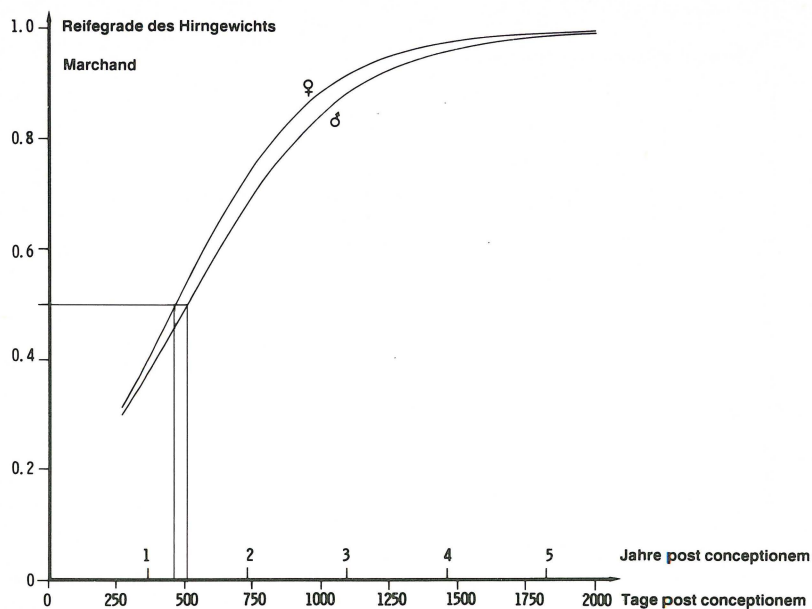


Abbildung 3 Die Reifegrade der logistischen Wachstumsfunktion der männlichen und weiblichen Hirngewichte wurden aus der Stichprobe von MARCHAND (1902) berechnet.

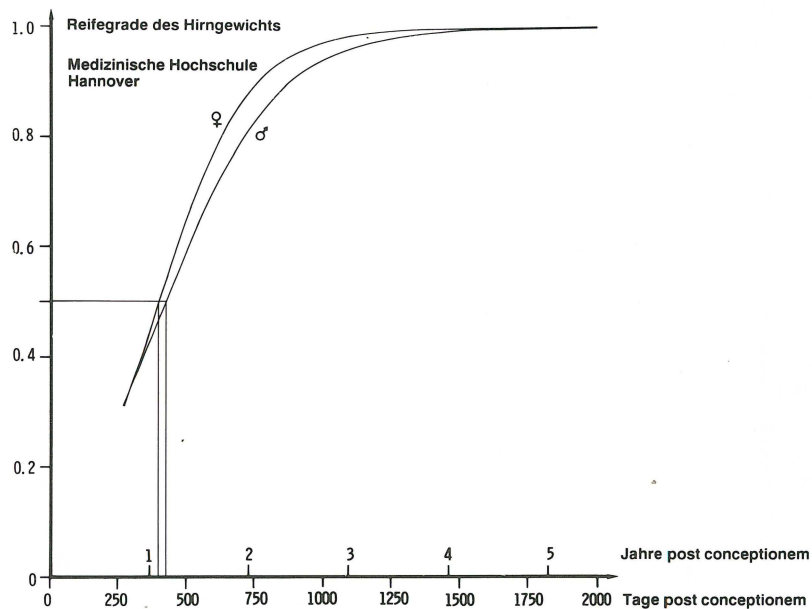


Abbildung 4 Die Reifegrade der logistischen Wachstumsfunktion der männlichen und weiblichen Hirngewichte wurden aus der Stichprobe der Medizinischen Hochschule Hannover ermittelt.

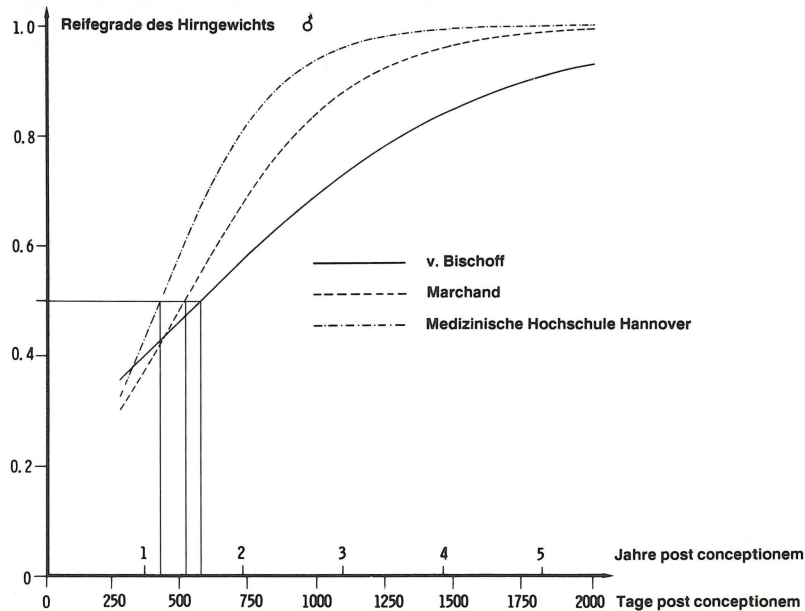


Abbildung 5 Die Reifegrade der logistischen Wachstumsfunktion männlicher Hirngewichte wurden von 3 Stichproben verglichen: v. BISCHOFF aus den Jahren vor 1880, MARCHAND aus den Jahren vor 1902 und Medizinische Hochschule Hannover aus den Jahren 1974–1978.

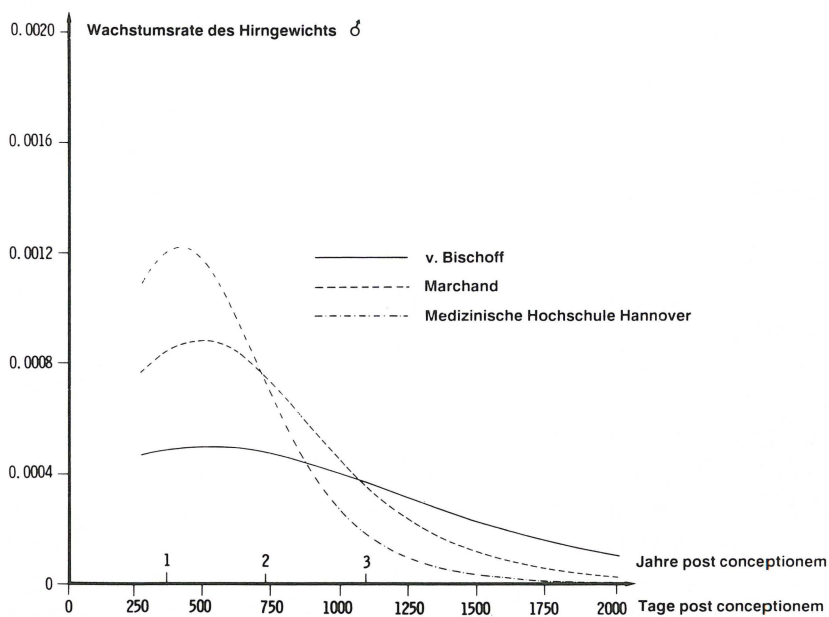


Abbildung 6 Die Wachstumsraten der Reifegrade männlicher Hirngewichte der 3 Stichproben der Abbildung 5.

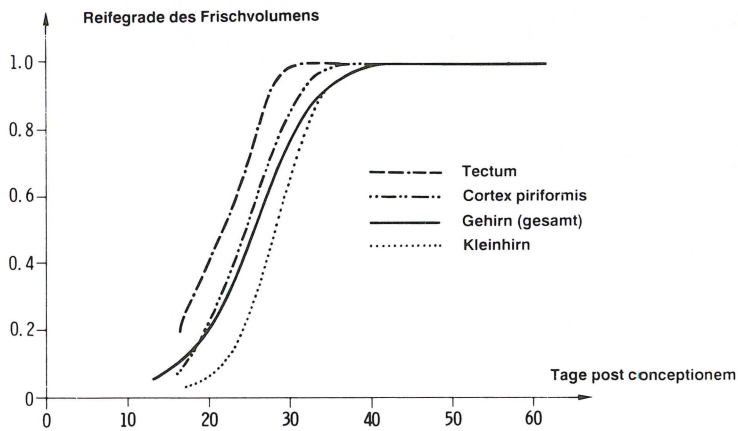


Abbildung 7 Die Reifegrade der logistischen Wachstumsfunktion des Frischvolumens des Tectum, des Cortex piriformis, des Gesamthirns und des Kleinhirns der Albinomaus (KRETSCHMANN, WINGERT 1971).

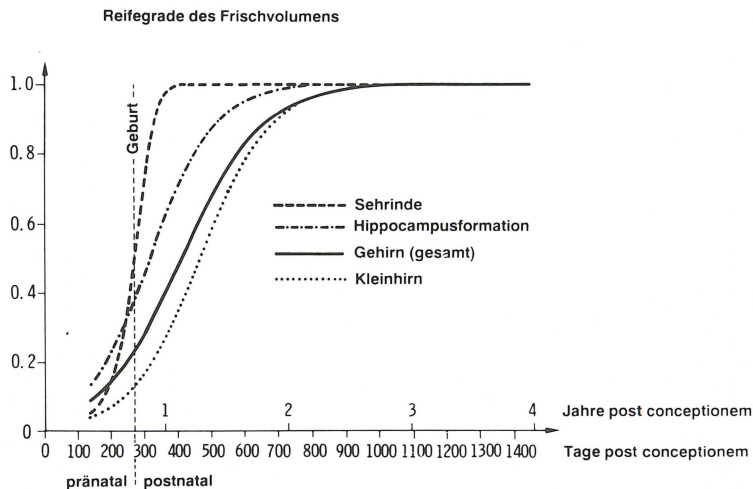


Abbildung 8 Die Reifegrade der logistischen Wachstumsfunktion des Frischvolumens der Sehrinde, der Hippocampusformation, des Gesamthirns und des Kleinhirns vom Menschen (KRETSCHMANN et al. 1986).

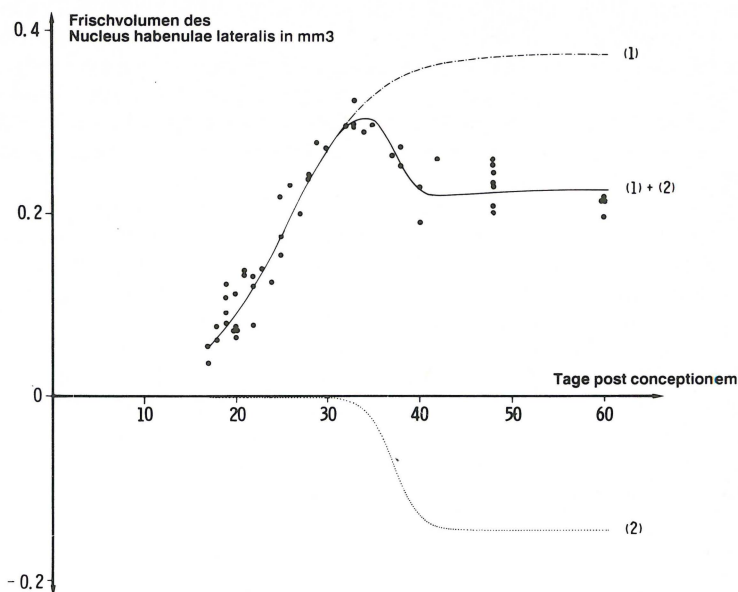


Abbildung 9 Das Frischvolumen des Nucleus habenulae lateralis der Albinomaus wurde gegen das Lebensalter aufgetragen und mit der logistischen Mehrkomponentenanalyse Komb bestimmt (ZILLES et al. 1976 a).

1986 b, RIEDEL et. al. 1989, SAUER et al. 1983, WINGERT 1969 a, ZILLES 1978, ZILLES et al. 1976 a, 1976 b, 1976 c). Diese Untersuchungen mit parametrischen Funktionen basieren auf Querschnittsanalysen, weil die Urdaten postmortal von Gehirnen gewonnen wurden. Für Längsschnittstudien bewährten sich neuerdings Kern-Schätzungen (GASSER et al. 1984, 1985). Übersichten der mathematischen Grundlagen bei Wachstumsvorgängen finden sich bei FELDMANN (1979), GOLDSTEIN (1978), MARUBINI (1978), PEIL (1978) und SCHARF (1977, 1981). Mit Hilfe des Programms Logi wurden die postnatalen Hirngewichte berechnet, die vor etwa 90 Jahren von MARCHAND an über 700 Gehirnen des männlichen und an über 400 Gehirnen des weiblichen Geschlechts bestimmt worden waren. Die Kurven zeigen, daß das Maximum der Kurven im männlichen Geschlecht höher als im weiblichen Geschlecht liegt. Der Parameter P_1 ist im männlichen Geschlecht etwa 130 g größer als im weiblichen Geschlecht (Abb. 1 a, b). Die Stichprobe aus den Jahren 1974–78 zeigt ähnliche Unterschiede (Abb. 2 a, b). Diese Ergebnisse bestätigen bekannte Fakten.

Die Kurven der Reifegrade zeigen ein in der Neuroanatomie bisher unbekanntes Phänomen. Es konnte nachgewiesen werden, daß im frühkindlichen Alter die entsprechenden Reifegrade im weiblichen Geschlecht früher als im männlichen Geschlecht erreicht werden (Abb. 3, 4). Diese quantitativen Ergebnisse korrelieren mit der Erfahrung, daß die geistige Reifung der Mädchen im Mittel schneller als die der Jungen verläuft. Ein Vergleich der Reifegrade der Hirnentwicklung aus verschiedenen Zeitperioden zeigt säkuläre Unterschiede: Heute verläuft die Hirnentwicklung schneller als vor 90, bzw. 110 Jahren (Abb. 5). Bei den australischen Ureinwohnern, den Aborigines, verläuft heute noch die postnatale Hirnentwicklung langsamer als bei den Kaukasiern (RIEDEL et al. 1989).

Wendet man die logistische Wachstumsfunktion auf einzelne Hirnregionen an, so läßt sich nachweisen, daß sich die Hirnregionen unterschiedlich, d. h. heterogen entwickeln. Es gibt

unter den Hirnregionen Früh- und Spätentwickler. Dies trifft bei einem Säuger wie der Albino Maus, aber auch für die menschliche Hirnentwicklung zu (Abb. 7, 8). Diesem kurzen Blick in die neurobiologischen Wachstumsprozesse soll ein Phänomen angefügt werden, das bei diesen Studien erstmals quantitativ nachgewiesen wurde. Es zeigte sich, daß in der frühen postnatalen Phase Nervengewebe im Überschuß gebildet werden kann, das dann später wieder abgebaut wird (Abb. 9). Wahrscheinlich wählt die Evolution auf diese Weise die neuronalen Anschlüsse aus, die eine bestmögliche synaptische Vernetzung der Nervenzellen ermöglichen.

Im ganzen gesehen zeigen diese Wachstumsanalysen zeitliche Ordnungen der Wachstumsverläufe. Diese Wachstumskurven von gesunden Gehirnen bilden die Grundlage, um krankhafte Hirnentwicklungen abgrenzen zu können. Außerdem lenken sie unsere Aufmerksamkeit auf generationsübergreifende Probleme, wie an der säkulären Veränderung der Hirnentwicklung demonstriert wurde. Die wissenschaftlichen Leistungen FRIEDRICH WINGERTS sind für die Wachstumsfunktion grundlegend und bleiben Vorbild. Welche Bedeutung diesen Vorgängen heute beigemessen werden muß, erkennt man an den Schriften des Club of Rome, die sich mit den Grenzen des Wachstums beschäftigen. Es müssen Konsequenzen gezogen werden, wenn es um das Leben der kommenden Generationen geht. Diese Verantwortung der Wissenschaft hat FRIEDRICH WINGERT betont. Wir haben allen Grund, seiner großen Leistungen zu gedenken.

Literatur

- BERKSON, J. (1953): A statistically precise and relatively simple method of estimating the bioassay with quantal response, based on the logistic function. *Amer. statist. Association J.* **48**, 565–599.
- BISCHOFF, L. W. von (1880): *Das Hirngewicht des Menschen*. Neusser, Bonn.
- FELDMANN, U. (1979): *Wachstumskinetik*. Springer, Berlin, Heidelberg, New York
- GASSER, T., KÖHLER, W., MÜLLER, H. G., KNEIP, A., LARGO, R., MOLINARI, L., PRADER, A. (1984): Velocity and acceleration of height growth using kernel estimation. *Annals of Human Biology* **11**, 397–411.
- GASSER, T., MÜLLER, H. G., KÖHLER, W., PRADER, A., LARGO, R., MOLINARI, L. (1985): An analysis of the mid-growth spurt and of the adolescent growth spurt based on acceleration. *Annals of Human Biology* **12**, 129–148.
- GOLDSTEIN, H. (1978): Sampling for growth studies. In: FALKNER, F. and TANNER, J. M. (Eds.) *Human growth*. Plenum Press, New York and London, **1**, 183–208.
- KRETSCHMANN, H.-J., WINGERT, F. (1971): *Computeranwendungen bei Wachstumsproblemen in Biologie und Medizin*. Springer, Berlin, Heidelberg, New York.
- KRETSCHMANN, H.-J., SCHLEICHER, A., WINGERT, F., ZILLES, K., LÖBLICH, H.-J. (1979): Human brain growth in the 19th and 20th century. *J. neurol. Sci.* **40**, 169–188.
- KRETSCHMANN, H.-J., KAMMRADT, G., KRAUTHAUSEN, I., SAUER, B., WINGERT, F. (1986 a): Brain growth in man. *Bibliotheca anatomica* (Basel: Karger) **28**, 1–26.
- KRETSCHMANN, H.-J., KAMMRADT, G., KRAUTHAUSEN, I., SAUER, B., WINGERT, F. (1986 b): Growth of the hippocampal formation in man. *Bibliotheca anatomica* (Basel: Karger) **28**, 27–52.
- MARUBINI, E. (1978): Mathematical handling of long-term longitudinal data. In: FALKNER, F. and TANNER, J. M. (Eds.): *Human growth*. Plenum Press, New York and London, 209–225.
- PEARL, R. REED, L. J. (1923): Skew-growth curves. *Proc. Nat. Acad. Sci., Washington* **11**, 16–22.
- PEIL, J. (1978): Das Logistische Wachstums-gesetz und seine Erweiterungen. *Gegenbaurs morph. Jb.* **124**, 524–545.
- RIEDEL, A., KLEKAMP, J., HARPER, C., KRETSCHMANN, H.-J. (1989): Morphometric study on the postnatal growth of the cerebellum of Australian aborigines and Caucasians. *Brain growth* **499**, 333–343.
- SAUER, B., KAMMRADT, G., KRAUTHAUSEN, I., KRETSCHMANN, H.-J., LANGE, H. W., WINGERT, F. (1983): Qualitative und quantitative development of the visual cortex in man. *J. comp. Neurol.* **214**, 441–450.
- SCHARF, J.-H. (1977): Wachstum. *Verh. anat. Ges.* **71**, 29–58.
- SCHARF, J.-H. (1981): Möglichkeiten der mathematischen Formulierung von Wachstumsprozessen. *Gegenbaurs morph. Jb., Lpzg.* **127**, 706–740.

- VERHULST, P. F. (1845): Recherches mathématique sur la loi d'accroissement de la population. Nouv. Mem. Acad. Roy. Sci. Belleslett., Bruxelles **18**, 1–38.
- VERHULST, P. F. (1847): Deuxième mémoire sur la loi d'accroissement de la population. Nouv. Mem. Acad. Roy. Sci. Belleslett., Bruxelles **20**, 1–32.
- WINGERT, F. (1969 a): Biometrische Analyse der Wachstumsfunktionen von Hirnteilen und Körpergewicht der Albinomäus. J. Hirnforsch. **11**, 133–197.
- WINGERT, F. (1969 b): Eine Verallgemeinerung der logistischen Wachstumsfunktion. Schriftenreihe des Deutschen Rechenzentrums Hft. S–9. Deutsches Rechenzentrum, Darmstadt.
- WINGERT, F. (1971): Eine Verallgemeinerung der logistischen Wachstumsfunktion. Biometrische Zeitschrift **13**, 33–78.
- ZILLES, K. (1978): Ontogenesis of the visual system. Advanc. Anat. Embryol. Cell Biol. **54**, 1–38.
- ZILLES, K., SCHLEICHER, A., WINGERT, F. (1976 a): Quantitative Analyse des Wachstums der Frischvolumina limbischer Kerngebiete im Diencephalon und Mesencephalon einer ontogenetischen Reihe von Albinomäusen. I. Nucleus habenulare. J. Hirnforsch. **17**, 1–10.
- ZILLES, K., SCHLEICHER, A., WINGERT, F. (1976 b): Quantitative Analyse des Wachstums der Frischvolumina limbischer Kerngebiete im Diencephalon und Mesencephalon einer ontogenetischen Reihe von Albinomäusen. II. Corpus mamillare. J. Hirnforsch. **17**, 11–20.
- ZILLES, K., SCHLEICHER, A., WINGERT, F. (1976 c): Quantitative Analyse des Wachstums der Frischvolumina limbischer Kerngebiete im Diencephalon und Mesencephalon einer ontogenetischen Reihe von Albinomäusen. III. Nucleus interpeduncularis. J. Hirnforsch. **17**, 21–29.

Anschrift des Verfassers: Prof. Dr. H.-J. Kretschmann, Medizinische Hochschule Hannover, Abteilung Neuroanatomie, Postfach 61 01 80, D-3000 Hannover 61

Biomathematik in Unterricht und Forschung

B. Schneider

Zusammenfassung

Die Entwicklung der Lernzielkataloge „Biomathematik“ und „Medizinische Statistik und Informationsverarbeitung“ werden geschildert und dabei besonders die Beiträge von FRIEDRICH WINGERT und seinen Mitarbeitern hervorgehoben.

Summary

The important contributions of FRIEDRICH WINGERT to the development of lectures in „biomathematics“ and „medical statistics and information processing“ are presented.

Biomathematik in Unterricht und Forschung

1960 hat der Wissenschaftsrat in seinen „Empfehlungen zum Ausbau der wissenschaftlichen Einrichtungen“ (Teil I. Wissenschaftliche Hochschulen. Nov. 1960) auch die Einrichtung eines neuen Fachgebiets empfohlen, das „Medizinische Statistik und Dokumentation“ genannt wurde. An den Empfehlungen hat maßgeblich SIEGFRIED KOLLER mitgearbeitet. KOLLER hat 1961 „Ein Institutsprojekt nach den Empfehlungen des Wissenschaftsrates zur Einrichtung von Lehrstühlen und Instituten für medizinische Statistik und Dokumentation“ veröffentlicht (KOLLER, 1961) und 1963 den ersten Lehrstuhl und das entsprechende Institut in Mainz übernommen. Seinen Bemühungen ist es im wesentlichen auch zu danken, daß bereits zu Beginn der 70er Jahre an den meisten Medizinischen Fakultäten und Hochschulen der Bundesrepublik Deutschland und Westberlins Lehrstühle und Institute für das neue Fachgebiet errichtet worden waren.

Nachdem das Fachgebiet etabliert war, stellte sich das Problem, die Lehrveranstaltungen zu diesem Fach zu systematisieren. Einer solchen Systematisierung kam die Neuordnung der Approbationsordnung und die Zentralisierung der Prüfungsfragen im Institut für Medizinische und Pharmazeutische Prüfungsfragen (IMP) entgegen. Für die einzelnen Fächer mußten Gegenstandskataloge erarbeitet und für die aufgeführten Gegenstände Multiple-choice-Fragen entworfen werden. Eine Kommission, die der leider zu früh verstorbene Kollege, EDWARD WALTER aus Freiburg leitete, hatte sich dieser Aufgabe angenommen und einen ersten Gegenstandskatalog für die Vorlesung und Übungen „Biomathematik“ erarbeitet. WALTER hatte 1974 erstmals auch Erläuterungen zu diesem Gegenstandskatalog veröffentlicht. 1976 wurde dieser Katalog bereits unter der Mitwirkung von FRIEDRICH WINGERT überarbeitet und ist seitdem gültig. Er unterteilt das Fachgebiet Biomathematik in 5 Abschnitte:

1. Mathematische Grundlagen (dieser Abschnitt wurde in der überarbeiteten Fassung in den Gegenstandskatalog der Physik verlegt).
2. Beschreibende Statistik.
3. Wahrscheinlichkeitsrechnung.
4. Statistische Methoden, Versuchsplanung.
5. Medizinische Bibliographie.

Hier sind einige Bemerkungen zur Bezeichnung „Biomathematik“ am Platze:

In einer 1959 posthum erschienenen Arbeit schrieb der Heidelberger Zoologe und Biometriker WILHELM LUDWIG, der als erster zum Sekretär (Schriftführer) der Deutschen Region der Biometrischen Gesellschaft (die damals noch gesamtdeutsche Region war) gewählt wurde und 1959 während des 6. Biometrischen Kolloquiums in Leipzig verstarb, „Über die Anfänge der Statistik und Biometrik“: „Eine für die Gegenwart gültige Definition der „Biometrik“ würde wohl, in Anlehnung an eine Textstelle auf einer der Umschlagseiten der Biometrics-Hefte, lauten müssen: Wissenschaft (oder Hilfswissenschaft) a) von der sparsamen Planung von Erfassungen und Experimenten und der erschöpfenden Auswertung der Ergebnisse (= Biometrik im klassischen Sinn) sowie b) von der Aufstellung quantitativer Hypothesen zur Erklärung biologischer Tatbestände und Vorgänge (oft kurz „Biomathematik“ genannt)“ LUDWIG (1959).

Zur Bezeichnung „Biometrik“ bemerkt W. LUDWIG: „In SCHUBERTS Handbuch (1835 ff.) werden zur Rechtfertigung des (hier nicht aus dem Italienischen übernommenen) Wortes Statistik, seiner Endung -ik, Diplomatie, Heraldik, Heuristik, Sphragistik (Siegelkunde) aufgeführt. Dieses leitet über zur Frage: Biometrie oder Biometrik? Die Griechen sprachen von „Ta mathematika“. Nach einem Kommentar des ARISTOTELES hing über dem Eingang von PLATOS Studier- und Unterrichtszimmer ein Schild, kein „ageometrikos“ möge seine Schwelle überschreiten. Schreibt man rasch die geläufigen Bezeichnungen von Wissenschaften hin, so findet man in der Mehrzahl die Endung -ik. Die meisten aus dem Klassischen stammenden und fast alle neueren Bezeichnungen für Disziplinen endigen so. Deshalb dürfte sich der Name „Biometrik“ empfehlen, zumal er auch im angelsächsischen Schrifttum üblich ist: die Zeitschrift „Biometrika“ (K. PEARSON), 1. Band 1901; „Biometrics“, 1. Band 1945“ LUDWIG (1959).

In diesem zuletzt genannten Punkt hat sich die Ansicht von W. LUDWIG nicht durchgesetzt. Überwiegend wird im deutschsprachigen Raum von „Biometrie“ geredet, während im Angelsächsischen sowohl die Bezeichnung „Biometrics“ als auch „Biometry“ vorkommen. Die Unterteilung der Biometrie in die beiden Teilbereiche, die LUDWIG „Biometrik im engeren Sinne“ und „Biomathematik“ genannt hatte und die heute meist als „Biostatistik“ und „Biomathematik“ bezeichnet werden, hat sich aber durchgesetzt.

Unter „Biomathematik“ werden vor allem „quantitative Theorien“ (so in der Satzung der Internationalen Biometrischen Gesellschaft) oder mathematische Modelle zur Beschreibung oder Erklärung biologischer Vorgänge verstanden. Diese Modelle können entweder deterministisch formuliert werden, etwa durch Differentialgleichungen (z. B. die Kompartimentmodelle in der Pharmakokinetik), oder durch Integrodifferentialgleichungen (z. B. die Lotka-Volterra'schen Gleichungen für das Wachstum von konkurrierenden Populationen) oder auch als stochastische Prozesse dargestellt werden. Die frühen Arbeiten von FRIEDRICH WINGERT über Wachstumsvorgänge, die er meist zusammen mit KRETSCHMANN durchgeführt und veröffentlicht hat, gehören weitgehend zu diesem Teilgebiet.

Für die im Gegenstandskatalog des Fachs Biomathematik aufgeführten Lehrinhalte ist aber nach dem heutigen Verständnis die Bezeichnung Biomathematik nicht zutreffend. Vielmehr gehören diese Inhalte fast ausschließlich (abgesehen von dem Gebiet der Wahrscheinlichkeitsrechnung und der Medizinischen Bibliographie) in das Gebiet der Biostatistik oder Medizinischen Statistik. In der Neufassung des Gegenstandskatalogs, die unter Leitung von FELDMANN erstellt wurde, wurde hierzu von der Kommission folgendes festgestellt: „Namensgebung: Es herrschte Einvernehmen, daß die Bezeichnung „Biomathematik“

weder die Lehrinhalte des bisherigen Kataloges noch die des vorliegenden Entwurfes abdeckt. Es wurde vorgeschlagen, den Begriff „Biometrie“ als zutreffend und bereits etabliert zu verwenden. Dies setzt eine entsprechende Änderung der Approbationsordnung voraus.“ Ich möchte allerdings im folgenden die Bezeichnung Biomathematik beibehalten.

Als FRIEDRICH WINGERT 1973 den Lehrstuhl für Medizinische Informatik und Dokumentation und die Leitung des Instituts für Medizinische Informatik und Biomathematik der Westfälischen Wilhelms-Universität Münster übernahm, hat er sich sofort zusammen mit seinen Mitarbeitern – von denen NIENHAUS, HEINECKE und HULTSCH besonders zu nennen sind – um die Ausgestaltung der Vorlesung und Übungen in Biomathematik intensiv gekümmert. Er hat in der Fachvertreterkommission für den Gegenstandskatalog mitgearbeitet und die zweite Fassung von 1976 entscheidend mitbestimmt. In Münster wurde bereits 1974 unter seiner Verantwortung eine Aufgabensammlung für Biomathematik erstellt, die Grundlage der praktischen Übungen für Biomathematik im Wintersemester 1974/75 und Sommersemester 1975 war. Diese wurde im März 1976 überarbeitet und kann als vorbildlich für die Veranstaltung angesehen werden. Zusammen mit seinen Fachkollegen in Nordrhein-Westfalen hat er sich an dem „Kollegium Biomathematik NW“: „Biomathematik für Mediziner“ beteiligt, das 1975 in erster Auflage und bereits 1976 in zweiter Auflage beim Springer-Verlag erschien. Schließlich sind vor allem noch seine Beiträge zur Entwicklung von Multiple-choice-Prüfungsfragen zu nennen. Eine Sammlung dieser Fragen erschien unter dem Titel „Examensfragen Biomathematik“ unter der Mitautorenschaft von HEINECKE, A., HULTSCH, E. und REPGES, R. ebenfalls 1975 im Springer Verlag. Ich benutze seit Jahren diese Multiple-Choice-Fragen in den Klausuren, die ich zur Überprüfung des „Erfolgs“ der Vorlesungen und Übungen abhalte, und bewundere dabei die Praxisbezogenheit der Fragen und die Variationsmöglichkeiten der Fragetypen (z. B. Typ A, der nur eine einfache Frage beinhaltet, auf die die richtige Antwort gefunden werden muß, oder Typ C, bei dem eine logische Folgerung vorgegeben ist, wobei die Richtigkeit der Prämissen und der Verknüpfung beurteilt werden soll) und der vorgeschlagenen Antwortmöglichkeiten, von denen ja nur eine einzige richtig sein darf. Einige Beispiele dieser Multiple-choice-Fragen aus dem Arbeitskreis Münster habe ich im folgenden zusammengestellt (Abb. 1).

Ich komme zurück auf die praktischen Übungen in Biomathematik, die in Münster 1974 und 1976 erarbeitet wurden. Vorbildlich an diesen Übungen ist, daß sie nicht von der mathematischen Methode, sondern von der praktischen Anwendung ausgehen. Zu Beginn wird eine praktische Aufgabe gegeben. Die dabei (real oder fiktiv) erfaßten Daten werden durchgehend benutzt, um – gewissermaßen zwangsläufig – die Anwendung der verschiedenen biometrischen Methoden zu demonstrieren bzw. auf Fehlanwendungen hinzuweisen. In der ersten Auflage wird zur Gewinnung von Übungsbeispielen ein Fragebogen vorgelegt, in den die Studentinnen und Studenten ihre persönlichen Daten (Geschlecht, Alter, Größe, Gewicht, Augenfarbe, Haarfarbe) eintragen sollen. Diese Daten werden später zur Einübung der verschiedenen Methoden benutzt. Dieses Verfahren hat den Vorteil, daß von Anfang an die Studentinnen und Studenten aktiv zur Mitarbeit aufgefordert werden. Es hat sich sehr bewährt, und ich mache auch immer wieder Gebrauch von diesem Verfahren. In der zweiten Auflage aus dem Jahr 1976, die schon im Titelumschlag eine gewisse Originalität aufweist, wird ein ophthalmologisches Problem an den Beginn gestellt. Analog zum originellen Vorgehen von IMMICH (1971) wird eine fiktive Pressemitteilung über einen klinischen Großversuch mit „Miraculin“ vorgelegt (Abb. 2), in der ein neues Präparat bei akutem Glaukomanfall angepriesen wird. Es entspricht der Genauigkeit von WINGERT und ist vielleicht auch auf seine langjährige Zusammenarbeit mit KRETSCHMANN zurückzuführen, daß er auch noch die topographische Anatomie des Auges beifügt, um die später eingeführten Meßgrößen, wie z. B. Kammerwinkel, Augeninnendruck u. ä., den Studenten anschaulich vorführen zu können. Die Vorbereitung einer korrekten, kontrollierten Studie zum Vergleich des Wundermittels „Miraculin“ mit der Standardtherapie „Ocultan“ wird

MS - 2 - 20 - 1 (++)

Fragetyp D

Welche der folgenden Zusammenhänge sind determiniert, welche sind zufällig?

- 1 Anzahl der Würfe mit einem Würfel und Summe der gewürfelten Augenzahlen,
- 2 Fallhöhe und Aufprallgeschwindigkeit eines Körpers,
- 3 Masse und Gewicht eines Körpers,
- 4 zur Heilung einer Krankheit angewandte Therapie und Heilerfolg,
- 5 Tageszeit und Stand der Sonne zu diesem Zeitpunkt,
- 6 tatsächliche Krankheit eines Patienten und Diagnose des Arztes.

Wählen Sie bitte unter den folgenden Aussagekombinationen diejenige, die Sie für zutreffend halten:

Zufällig sind nur die Zusammenhänge

- | | |
|---------------|---------------|
| A 1, | D 1, 2 und 5, |
| B 1, 4 und 6, | E 1 und 6. |
| C 1, 4 und 5, | |

MS - 4 - 84 - 2 (++)

Fragetyp A₁

Der KLINISCHE DIAGNOSENSCHLÜSSEL

- A ist ein sechstelliger numerischer Schlüssel,
- B unterscheidet die semantischen Kategorien TOPOGRAPHIE und FUNKTIONSTÖRUNGEN,
- C unterscheidet die semantischen Kategorien TOPOGRAPHIE und ÄTIOLOGIE,
- D unterscheidet die semantischen Kategorien TOPOGRAPHIE und PATHISCHER PROZESS,
- E ist mit der INTERNATIONAL CLASSIFICATION OF DISEASES identisch.

MS - 1 - 7,8 - 3 (+++)

Fragetyp C

Der Fettgehalt der Milch ist kein diskretes Merkmal
denn

der Fettgehalt der Milch ist kein qualitatives Merkmal.

Bitte kreuzen Sie die Antwort A - E an, die nach Ihrer Meinung die beiden Feststellungen und ihre Verknüpfung richtig beurteilt:

Antwort	Feststellung 1	Feststellung 2	Verknüpfung
A	richtig	richtig	richtig
B	richtig	richtig	falsch
C	richtig	falsch	-
D	falsch	richtig	-
E	falsch	falsch	-

MS - 4 - 85 - 2 (+)

Fragetyp A₁

MEDLARS und DIMDI sind Bezeichnungen für

- A Gruppen von Pharmaka,
- B Namen von Krankenhausinformationssystemen,
- C Literaturdienste,
- D klinische Diagnoseschlüssel,
- E Handlochkartensysteme.

Abbildung 1: Beispiele von Multiple-choice-Prüfungsfragen aus Münster

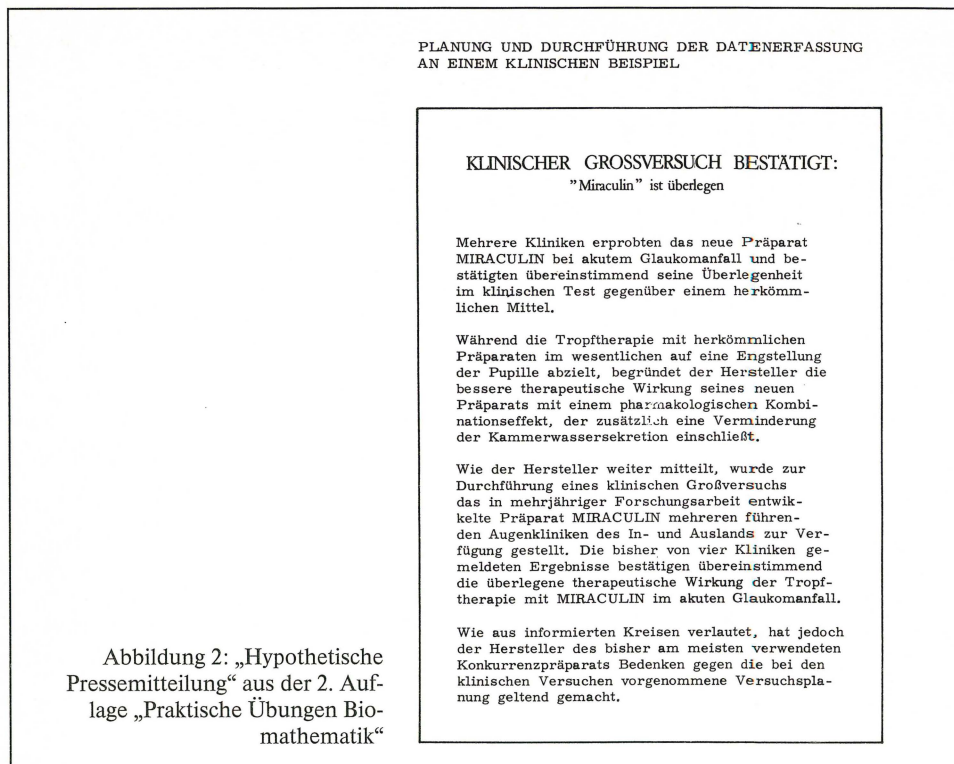


Abbildung 2: „Hypothetische Pressemitteilung“ aus der 2. Auflage „Praktische Übungen Biomathematik“

anhand eines Erfassungsbogens und der dazugehörigen Lochkarte demonstriert (Abb. 3). Hypothetische Daten für diese kontrollierte Studie werden dazu benutzt, um z. B. Zusammenhangsanalysen (Regression, Korrelation) zu demonstrieren oder die Binomialverteilung als Prüfverteilung des Vorzeichentests zur Überprüfung der Nullhypothese, daß in einem intraindividuellen Vergleich die Wahrscheinlichkeit für eine bessere Wirkung von „Miraculin“ als von „Ocultan“ $1/2$ ist, anschaulich darzustellen (Abb. 4).

Dieser kurze Gang durch die praktischen Übungen aus dem Jahr 1976 zeigt, wie praxisnah und auch modern diese Übungen schon damals angelegt waren. Es gehört z. B. zu den Grundsätzen bei der in den letzten Jahren erfolgten Neubearbeitung des Gegenstandskatalogs „die Anwendungsrelevanz des Fachs Biometrie für die Medizin, insbesondere für die klinische Medizin, besser herauszuarbeiten und deutlicher darzustellen“. „Der zu lehrende mathematische Formalismus sollte auf das für Studierende der Medizin unbedingt notwendige Maß reduziert und die Aufgabe des Fachs „Biometrie“, die in der Bereitstellung und Anwendung von Methoden zur Erkenntnisgewinnung in der Medizin besteht, hervorgehoben werden.“ Die Gestaltung des Unterrichts, wie sie FRIEDRICH WINGERT in Münster eingeführt hat, kann hierfür Vorbild sein.

Die Aktivitäten FRIEDRICH WINGERTS zur Verankerung des Fachgebiets in der Approbations- und Prüfungsordnung für Mediziner gingen aber noch weiter. Von Anfang an hat er sich in der Kommission der Fachvertreter mit dafür eingesetzt, auch über die elementaren statistischen Verfahren hinausgehende Verfahren und insbesondere die Verfahren der Medizinischen Informatik in der Prüfungsordnung und der Approbationsordnung zu verankern. Die Bezeichnung „Medizinische Informatik“ hat inzwischen die ursprüngliche Bezeichnung „Dokumentation“ abgelöst. Diese Bemühungen waren erfolgreich. Ab 1978 existiert der Gegenstandskatalog 3 für den zweiten Abschnitt der ärztlichen Prüfung, der

- Klinische Anwendungen in der Statistik (kontrollierte klinische Therapiestudie, retrospektive Studien (Fallkontrollstudien), prospektive Studien (Kohortenstudien), Verfahren zur Unterstützung der ärztlichen Diagnostik.

- Medizinische Informatik (Grundbegriffe [Signale, Daten], Informationsalgorithmus, Dokumentation, Anwendungssysteme, Datenschutz).

Um vor allem den Aspekten der Medizinischen Informatik mehr Gewicht zu verschaffen, hat er ein eigenes Lehrbuch zu diesem Fach geschrieben, das 1979 in deutscher Sprache (im

Klinischer Vergleich der Medikamente Miraculin und Ocultan			
Daten von Patienten mit Glaucoma acutum congestivum			
1. Laufende Nummer			01
2. Tropftherapie (Ocultan = 1, Miraculin = 2)			3
3. Krankenblattnummer		093568	
4. Aufnahme datum	10	181168	
5. Geburtsdatum	16	150394	
6. Geschlecht (männlich = 1, weiblich = 2)			2

7. Körpergröße in cm	23	168	
8. Körpergewicht in kg	26	075	
9. {Blutdruck (RR) in mm Hg} systolisch	29	225	
{ bei der Aufnahme } diastolisch	32	130	

11. Erfaßtes Auge (rechts = 1, links = 2)			3
12. Anzahl früherer Glaukomanfälle auf diesem Auge			3
13. Kammerwinkel des Auges (weit = 1, normal = 2, eng = 3)			3
14. Augeninnendruck bei der Aufnahme in mm Hg	38	071	
15. Visus bei der Aufnahme	41		

16. am 1. Tag	44	058	
17. Augeninnendruck unter { am 2. Tag	47	048	
Therapie in mm Hg { am 3. Tag	50	041	
(Messung jeweils 9.00 Uhr) { am 4. Tag	53	036	
20. am 5. Tag	56	038	

21. Anzahl der Behandlungstage bis zur Operation			9
22. Visus bei der Entlassung	60	040	

Abbildung 3: Erfassungsbogen und Lochkarten für eine klinische Prüfung „Miraculin“

[illegible]

Aufgabe 3.3: INTERPRETATION VON KORRELATIONEN

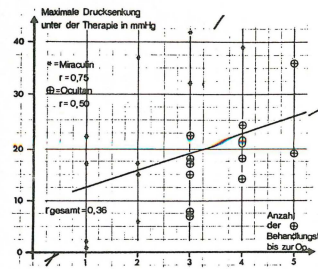


Bild 15: Maximale Drucksenkung unter der Therapie in Abhängigkeit von der Anzahl der Behandlungstage bis zur Operation.
Daten nach Tabelle 6

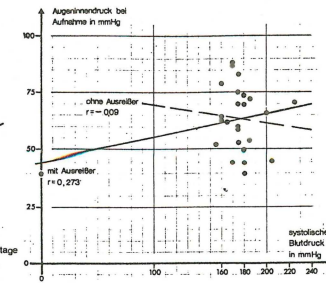


Bild 16: Augeninnendruck bei der Aufnahme in Abhängigkeit vom systolischen Blutdruck.
Daten nach Tabelle 6

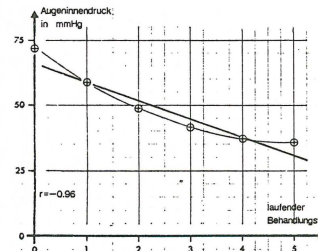


Bild 17: Augeninnendruck in Abhängigkeit vom laufenden Behandlungstag bei einem Patienten.
Daten nach Bild 3

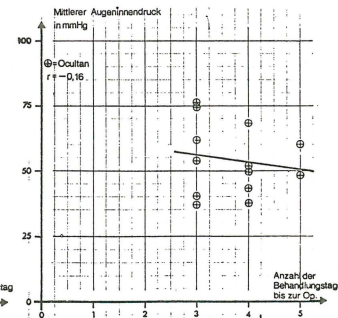


Bild 18: Mittlerer Augeninnendruck in Abhängigkeit von der Anzahl der Behandlungstage bis zur Operation bei Patienten unter Ocultan-Tropftherapie

Interpretieren Sie Bild 15 bis 18!
Welche dieser Regressionen sind sinnvoll?

$i =$	1	2	3	4	5	6
$A_1 =$			$e_8 = (OOO MM)$ $e_{13} = (MOO MO)$ $e_{14} = (OMO MO)$ $e_{15} = (OOM MO)$ $e_{21} = (MOO OM)$ $e_{16} = (OOO MO)$ $e_{24} = (OOO OM)$ $e_{29} = (MOO OO)$ $e_{30} = (OMO OO)$ $e_{32} = (OOO OO)$	$e_5 = (MOO MM)$ $e_6 = (OMO MM)$ $e_7 = (OOM MM)$ $e_{10} = (MMO MO)$ $e_{11} = (MOM MO)$ $e_{12} = (OMM MO)$ $e_{18} = (MMO OM)$ $e_{19} = (MOM OM)$ $e_{20} = (OMM OM)$ $e_{25} = (MMM OM)$ $e_{28} = (OMM OO)$	$e_2 = (MMO MM)$ $e_3 = (MOM MM)$ $e_4 = (OMM MM)$ $e_9 = (MMM MO)$ $e_{17} = (MMM OM)$	
$P(A_1) =$	1/32	5/32	10/32	10/32	5/32	1/32

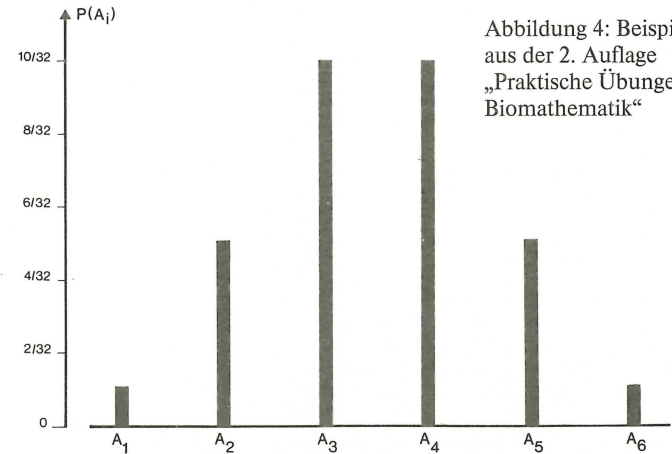


Abbildung 4: Beispiele aus der 2. Auflage „Praktische Übungen Biomathematik“

Bild 22: Tabelle und Blockdiagramm der Wahrscheinlichkeiten der Ereignisse $A_1 = \{\text{Miraculin wirkt bei (i-1) der 5 Patienten besser als Ocultan}\}$

Teubner-Verlag) und 1981 in englischer Sprache (im Springer-Verlag) erschienen ist. Es ist das erste und bisher einzige in Deutschland erschienene Lehrbuch der Medizinischen Informatik und dürfte auch im angelsächsischen Bereich kaum Konkurrenz haben. Im Vorwort zur deutschen Ausgabe weist er ausdrücklich darauf hin, daß die bisherige Situation „bei dem großen und breit gefächerten Bedarf an Unterstützung bei der Lösung medizinischer Probleme zu einem sehr inhomogenen Bild des Faches geführt“ hat, „das dem interessierten Anfänger einen Überblick erschwert“.

„Mit der zweiten Verordnung zur Änderung der Approbationsordnung für Ärzte (ÄAppO) vom 24. 2. 1978 wird eine Intensivierung der Ausbildung in „Biomathematik“/„Medizinische Statistik“/„Dokumentation“ gefordert.“ Um diesen Mängeln abzuweichen, hat er in dem Lehrbuch „Das Fach der Medizinischen Informatik“ definiert und die Wissensinhalte, die zu diesem Fach zählen, systematisch dargestellt. „Medizinische Informatik“ versteht er „als die Vereinigung aller wesentlichen Verfahren der Informatik in der Medizin und der Aufbereitung medizinischer Daten für den Einsatz solcher Verfahren“ WINGERT (1979). In dem Buch wird zunächst ein Überblick über die wichtigsten Grundbegriffe und Verfahren der Informatik gegeben. Ein weiteres Kapitel ist der Klassifikation gewidmet und ein weiteres der medizinischen Linguistik; zwei Spezialgebiete, die FRIEDRICH WINGERT besonders weiterentwickelt und denen er die letzten Jahre seines Lebens fast ausschließlich gewidmet hat. Als Teilbereiche der Medizinischen Informatik behandelt er:

- Datenmodelle, Datenstrukturen,
- Datenerfassung und Dateneingabe,
- Datenspeicherung,
- Informationsbildung,
- Informationswiedergabe,
- Datenintegrität.

Schließlich werden noch spezielle Anwendungen der Medizinischen Informatik, wie Prozeßdatenverarbeitung (Biosignalverarbeitung, Bildverarbeitung, Labordatenverarbeitung u. ä.), betriebswirtschaftliche Anwendungen und Informationssysteme (Kommunikationssysteme, Krankenhausinformationssysteme), Krankheitsregister und Bibliographien behandelt.

Zum Schluß stellt sich die Frage, die mir fast in jedem Jahr von den Studentinnen und Studenten gestellt wird: „Wozu benötigt der Arzt dieses Wissen?“

Zur Beantwortung dieser Frage führe ich meist zwei Argumente an:

- Der Arzt ist heute nicht mehr „Leibarzt“ einiger weniger, sondern für die Bevölkerung verantwortlich. Er darf sich daher nicht darauf beschränken, nur das Einzelbild, den einzelnen Patienten, zu beurteilen und zu bewerten, sondern muß auch in der Lage sein, den Gesundheitszustand der Population zu berücksichtigen und die Auswirkungen von Gesundheitsmaßnahmen auf diese Population zu beurteilen. Dies genau kann er aber nur mit Hilfe der biometrischen Methoden.
- Ein an der Universität ausgebildeter Arzt sollte auch in der Lage sein, medizinische und insbesondere klinisch-therapeutische Forschung zu betreiben. Jede Forschung zielt aber nicht nur auf den Einzelfall, sondern immer auf das „Allgemeine“. In der Empirie bedeutet dieses Allgemeine die Population. Deshalb ist eine medizinische und insbesondere klinisch-therapeutische Forschung ohne biometrische Methoden nicht möglich.

Diese letzte Auffassung wird durch die Fakten gestützt. Zumindest angesehene wissenschaftliche Zeitschriften akzeptieren keine Arbeiten – seien es experimentell-medizinische oder klinisch-therapeutische Arbeiten –, bei denen adäquate biometrische Versuchsplanung und -auswertung fehlen. Viele dieser Zeitschriften haben auch Biometriker in ihrem Redaktions- oder Referentenstab. Darüber hinaus wird in den seit Ende 1987 gültigen „Grundsätzen für die ordnungsgemäße Durchführung der klinischen Prüfung von Arzneimitteln“ des Bundesministeriums für Jugend, Familie, Frauen und Gesundheit bei der Planung, Durchführung, Auswertung und Beurteilung klinischer Arzneimittelprüfungen der

Einsatz biometrischer Methoden zwingend vorgeschrieben. So heißt es z. B. in diesen Richtlinien: „Biometrische Überlegungen sind so früh wie möglich anzustellen.“ „Der Leiter der klinischen Prüfung, der verantwortliche Biometriker und die durchführenden Ärzte müssen für die Durchführung der klinischen Prüfung qualifiziert sein.“ „Nach Abschluß der Prüfung ist ein Bericht zu erstellen, der eine biometrische Auswertung und eine Bewertung der Ergebnisse aus medizinischer Sicht enthält.“ Diese Argumente dürften ausreichend belegen, daß es sich bei dem Fach „Biomathematik“, an dessen Einführung und Ausgestaltung FRIEDRICH WINGERT entscheidend Anteil genommen hat, nicht um eine exotische Laune einiger weniger handelt, sondern um ein für die Ausbildung des Arztes notwendiges Fach, das darüber hinaus für die klinische Forschung unentbehrlich ist.

Literatur

- IMMICH, H. (1971): Therapie-Effekte und Statistik. Die Kapsel 27, 1051 – 1114
 KOLLER, S. (1961): Ein Institutsprojekt nach den Empfehlungen des Wissenschaftsrates zur Einrichtung von Lehrstühlen und Instituten für medizinische Statistik und Dokumentation. Medical Documentation 5, 29 – 35
 LUDWIG, W. (1959): Über die Anfänge der Statistik und Biometrik. Biometrische Zeitschrift 1, 71 – 80
 WINGERT, F. (1979): Medizinische Informatik. G. B. Teubner, Stuttgart

Anschrift des Verfassers: Prof. Dr. Berthold Schneider, Medizinische Hochschule Hannover, Institut für Biometrie, Konstanty-Gutschow-Straße 8, D-3000 Hannover 61

Indexing Medical Information: The Role of SNOMED

D. J. Rothwell

Preface

It is with great sadness that I come before you today to help celebrate the memory of a close friend and of a true scholar, Professor FRIEDRICH WINGERT. I had the privilege of working closely with Professor WINGERT during the last few years of his life. His was an enormously productive and important life – but one that was too short. We continue to feel the void created by his absence both as a person and as a professional colleague. Professor WINGERT possessed limitless intellectual energy – he was formal yet nevertheless a warm person. As I knew him, he was without pretension – he was as dedicated and hardworking a person as any I have known in my entire life. He demanded excellence of himself and from those around him – and largely received it. When with him, one had the feeling that they had to reach deep within themselves and to do their best – anything less was not acceptable. As we worked over the last few years of his life, I felt a special bond develop between us – one that reached beyond our professional interactions and touched each of us personally – one that even now is sorely missed.

His vitality and dedication are not with us today but his work is. Through his work he continues to live in each one of us. I would like to review with you today some elements of his work – to review the problem and issues to which he addressed himself with such a singleminded purpose. To review the current status of this work and to outline what remains to be done.

Introduction

Automated natural language systems which accept natural language data input, answer questions about a knowledge base, make inferences, and generate natural language responses are complex, and techniques for designing them are in the early stage of development. Much work is being done on different aspects of the systems to develop more complete grammars, lexicons, and knowledge representation schemes to broaden their domain and improve their efficiency and accessibility.

Natural language processing by computers requires that language data be organized into appropriate data structures which are capable of recording information at various levels. Levels include the lexical, morphological, syntactic, and contextual content of a statement to be processed.

This paper addresses the processing of natural language data in the restricted domain of diagnostic and therapeutic medical statements. In so doing, it avoids many of the problems encountered in free natural language processing and analysis. Furthermore, processing is focused on indexing, which is the most important intermediate objective of such systems.

The Systematized Nomenclature of Medicine (SNOMED) is offered as an example of a multidimensional nomenclature; it is a classification system for medical nomenclature. The structure of SNOMED is introduced with emphasis on its capabilities with respect to an artificial medical language.

Classification

For health statistics to be useful in providing comparison over time, expressed as trends or variations, widespread agreement must be reached regarding basic terminology, its definitions, and the structure of the classification employed by those who collect the data. At the present time, medical science is attempting to construct an indexing and classification system for its terminology, one which will serve as a vehicle for the efficient storage and retrieval of medical data.

From a theoretical viewpoint, classification acts as a method of displaying the relationships between subjects. Grouping animal and plant species into taxons and assigning them to families, orders, and phylla based on their morphologic similarities or some other agreed-upon criteria forms the basis of a biological classification. Even in biological systems, parallel classifications designed to serve differing but specific purposes exist for the same subject matter. Biological classifications have two major objectives: to serve as the basis for biological generalization and to serve as the key to an information storage and retrieval system.¹ It is still unclear whether a single classification can serve both purposes and perform satisfactorily in both.

Biological classifications are frequently hierarchal. Relationships between subjects can be shown while their distinctions are maintained, and previously differentiated subjects can be treated as a unit. This ranking, or hierarchy, within the classification is its greatest strength and without it generalization and data retrieval, a classification's two major objectives, cannot be achieved. The construction of these hierarchies, particularly ones designed to serve different purposes, is one of the principal challenges of the taxonomist. On the other hand, the number and kinds of descriptors needed for a hierarchal classification in a biological system and the assignment of relative weights to each descriptor are two of the more difficult design problems that face the taxonomist.

In biological systems, a classification often reflects a phylogenetic relationship among its members and, while this is not in itself a sufficient basis on which to establish a classification, it does serve as the core or anchor, the central theme about which the classification is built. For medical terminology and our attempt to build a classification of medical terms, there is no unanimity regarding the theme or core of a systematized vocabulary – that is, a classification of medical terminology suitable to a wide range of purposes. For the pathologist and many other medical specialties, anatomic structure in the form of topographic sites serves as the core around which a systematized vocabulary is formed.² This topographically-based format closely parallels phylogeny in a biological classification. Whether anatomic structure could fill a similar role for a comprehensive classification for all of medical terminology is as yet unknown. Determining the basis or organizing principle for a classification is a crucial part of designing the structure for a comprehensive classification of medical terminology.³

When applied to medical terminology, the word classification itself often means different things to its users. In published medical classification systems, the rationale by which the terms are divided into groups or classes is not stated explicitly, and consequently it is often difficult to determine. In other systems the division into classes is inconsistent or incongruous. Unless the basis of the classification system is specified, and unless there is agreement about which meaning of the word is intended, there will be confusion rather than clarity when describing the system as a classification.

In each of the decennial versions of the International Classification of Diseases (ICD),⁴ classification has been used in several ways. One is the "stuffing" of several different but distinct diseases into one rubric. In other words, multiple and presumably closely related conditions are assigned the same code number, resulting in loss of the ability to retrieve a discrete entity. In another sense, classification is interpreted to mean the basis for grouping all diseases into seventeen chapters, the headings of which are a mix of body systems, neoplasms, and etiologies such as infectious disease injury and poisoning. There is no attempt to use classification in the traditional sense of indicating the relationships that exist between more specific and less specific terms and conditions. This creates difficulties when attempting to retrieve groups of cases related to a specific topic. Furthermore, ICD as a classification violates one of the basic principles: one code for one concept.⁵ For example, all diabetes mellitus cases are coded to "Diseases of the Endocrine System", except for those which involve obstetrical patients – in which cases the diabetes is grouped with "Complications of Pregnancy, Childbirth and the Puerperium". These inconsistencies have led to the loss of information when retrieval is attempted, and to a serious loss of specificity within ICD.

Other classifications have been more rigid in establishing the basis for their systems. The Standard Nomenclature of Diseases and Operations (SNDO),⁶ originally published in 1929 and abandoned in 1959, was the first schema to recognize human anatomy as the core of a classification. A SNDO rubric consisted of two parts separated by a hyphen. To the left of the hyphen was a topographic code, in the right a number representing the disease entity. The disease code, such as that for inflammation, remained the same regardless of the site code to which it was attached. Thus the "inflammation" portion of the code was consistent whether it represented appendicitis (appendix – inflammation) or inflammation of another site. Since the coding of the disease was consistent, the classification could be built around a different axis, anatomic structure. All conditions relating to a topographic site could be found in one group of adjacent rubrics, just as inflammation at any or all sites could be identified.

Indexing Systems

At issue is the optimal way to represent the contents of a document, be it a medical record or some other source of medical-technical information, so that information can be retrieved and studied. Storage and retrieval of demographic information such as names, dates, sex and other fixed elements found in uniform hospital data information sets present no problem to the information scientist. Rather it is representation of the scientific information within the document or medical record which is the challenge.

The conventional view of the documentation process was described by HUTCHINS in 1967.⁷ It involves "the analysis of each document's content, a formulation of this content in a set of descriptors, and an organization of descriptors such that enquirors can match their search requests and not miss any documents relevant to that request".

The major feature of this indexing process is replacement of the natural language text by a set of greatly abbreviated descriptors. This is done not only because the decreased size of the resultant index is more cost effective, but also, and more importantly, because the descriptors bring out the essential features of the document. These descriptors serve as a summary of content emphasizing the document's salient points and relating them to other documents in a collection. The descriptors provide information of a more uniform kind than that found in the documents themselves. For this reason, the use of full document text as its own descriptors (which has been proposed by some information scientists) is undesirable even if it could be shown to be economical given the advanced state of data storage techniques available today.⁸

Perhaps the most successful and widely used indexing system in medicine today is the Index Medicus,⁹ a product of over one hundred years of continuous development at the National

Library of Medicine in Washington, D. C. The Index Medicus is a retrieval system based on a formal descriptor list or thesaurus. It offers access to relevant documents through a series of key descriptors linked to each other in a tree structure. It is, in other words, a controlled indexing vocabulary. The indexing and accessing of medical literature is accomplished by analysis of each document's relevant content using Medical Subject Headings (MESH), a thesaurus presently containing over 14,000 descriptive terms. MESH is organized into a hierarchy or tree structure consisting of fifteen categories and over a hundred subcategories covering all aspects of biomedicine. Each of the fifteen categories is the base of a tree containing up to eight levels of progressively more specific related terms. Each position in each tree has a unique numeric identifier associated with the descriptor term at that position. Synonyms and preferred terms are handled by extensive crossreferencing, and pointers are used to identify more specific and less specific terms related to a specific topic. Generally documents have numerous descriptors terms assigned to them, and complex concepts not found in the descriptor list can often be expressed by coordination of terms. For example, jejunitis may be expressed by the descriptors as jejunal disease and enteritis. Most descriptor terms are modified by one or more subterms which lend additional specificity and which provide direct linkage to the parent term.

Stability of MESH subject headings is maintained by yearly review and updating of terms. MESH represents the conventional technique of indexing and accessing medical literature. Though it is abundantly clear that MESH works, it is limited in its scope. MESH indexes topics found in the medical literature; it does not address the broad issues and topics of clinical medicine. The many favorable evaluations of the system have not determined whether it is in fact the best system to accomplish its stated purposes. Since it was designed for literature search only, it will not be known until it is expanded and tested whether a system of similar design would perform satisfactorily as a medical indexing system for each of the broad purposes we have outlined.

In the development of an indexing system it would be ideal if some idea of the kinds of requests made of the system were known. Its design would then be relatively easy. Designing a system of a general nature and expecting it to respond to unanticipated requests is far more difficult, yet this is what is expected of a good information storage and retrieval system. Often those who design indexing systems are not the same individuals who need to retrieve information from them. Frequently the designers make no attempt at serious retrieval in order to evaluate a system's strengths and shortcomings – a problem not unique to information science. Regardless of the type of system, requests must be well formulated for significant retrieval of information to be successful.

The first step in creating a medical information system is to establish rules for the inclusion of terms that are to be represented in the data base, i. e., the terminology or nomenclature. If the system is designed so that it will retrieve only those specific terms requested, the problem is straightforward and should not pose significant difficulty. If, on the other hand, the system is expected to show relationships between terms in the data base as well as to retrieve information concerning a broad area or topic, then the relations between terms must be specified. An indexing system or, in some form, a classification of terms must be constructed. However, the designers of such a system cannot expect a terminology to do the work of a retrieval system; that is only part of the process.¹⁰

The only real alternative to an indexing system of whatever design or merit is to store just the original data as natural language. Then just the original data can be retrieved. If any type of grouping is required for retrieval then the user is left with the task of enumerating all the possible variations of expressing a given concept in order to retrieve it, in addition superordinated concepts will not be apparent. If someone is searching data denoting inflammatory processes of the kidney, he has to identify all possible formulations of this fact. There is no doubt that the result of such a query are quite diverse among a sample of users.

A more intelligent retrieval system should support a user by supplying knowledge about

which elements are relevant for a query. But then the debate starts all over again. Labeling data according to concepts such as kidney or inflammation is nothing else than making explicit the projections of a point in an information space with regard to a predefined set of axes and of discrete points of these axes. The concepts and their values that are selected for such labels together with the relations between the concepts (kidney is part of the urinary tract) form the basis for a multidimensional nomenclature. The Systematized Nomenclature of Medicine (SNOMED)¹¹ is an example of such a nomenclature; its structure will be shown in the next section. Before proceeding however, additional features needed in a well developed indexing system will be explored.

The conclusion is that intelligent documentation systems cannot totally discard classifications. Moreover, the availability of at least one classification is a necessary condition for a good documentation system. Classifications are not a necessary evil but a very effective way of representing knowledge about a domain of discourse.^{12, 13}

Indexing and Coding

Nomenclature, indices and classifications have been created by various medical specialties and disciplines to address the information problems encountered by each and to satisfy their own requirements for retrieval.

Not only do the terminologies of these systems differ since they involve varying disciplines, but so do the structures of the classifications used by each. The differences in structure make translation of information from one system to another difficult and often impossible. Techniques developed to help overcome these difficulties and to bridge between independent information sources have been largely unsuccessful. There remains a continuing and increasing need to develop a coordinated, comprehensive and coherent indexing system for medical information.

Indexing is the processing of statements by applying transformations to a language expression which results in assignment to a set of descriptors or a target language (intermediate language). This is followed by encoding. Therefore, indexing includes much more than just encoding since actual intellectual achievements are required for the transformation. But, frequently, indexing and encoding are used as synonyms.

There are many different reasons for indexing and then representing this information with codes. A selection of these reasons is listed below.

Standardization of terminology:

- Coding requires a careful definition of terms.
- In retrieval the preferred term replaces the internal code.
- Multilingual or multidialectic terminology can be mapped onto the same representation.
- The code can be used as a terminal symbol of an intermediate language by different medical languages in order to facilitate translation.

Facilitation of data storage:

- Codes have usually a much more rigorous structure than language data. Therefore, data elements can be better defined by formal means and the construction of data aggregates is facilitated.
- Storage capacity requirements are reduced.
- Retrieval is facilitated, less expensive and more precise.

Representation of relations:

- Representation of relations is facilitated. This is especially true for synonymy and hyponymy (hierarchical coding). Therefore, retrieval is supported by the use of background knowledge represented in the code structure.

Coding may be done by human encoders knowing the terminology and the classification. If this approach is used then a common postulation is that coding is done by those who gen-

erate the data, i. e. physicians. There are some good reasons for this requirement but practice has often demonstrated that, in spite of theoretical advantages, the results are worse than when coding is performed by specially trained medical record officers.

The alternative is automated encoding which has some promising advantages:

Improvement of consistency:

- Algorithms neither get tired nor change their mind after some time. Therefore, the same input is always mapped onto the same code. This is even true for the errors. Random code assignment is excluded.

Improvement of correctness:

- Some automated systems produce better results than humans when coding large amounts of data.
- In some applications a semiautomated procedure may be the best solution, at least, as long as the system is not sufficiently stable. The automated encoder can be enhanced by implementing learning capabilities.
- Error checking can be formulated explicitly.

Increase in encoding adaptability:

- In a manual coding system it is usually impossible to recode a large data set when coding rules or classification change. In an automated system this can be done usually with little difficulty.

For technical reasons, frequently, numerical systems are preferred because they are especially well-suited for representing hyponymy (hierarchic coding, Figure 1). Hierarchic coding considers the following relations:

Synonymy (Quasi-) synonyms have the same code

Hyponymy The code of a parent is the prefix of the codes of all siblings.

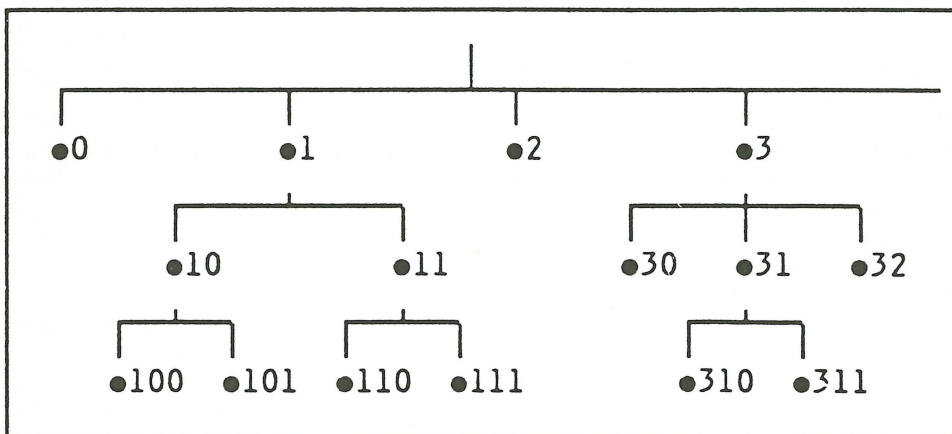
Cohyponymy Identical code prefixes.

Therefore, hierarchic coding represents more than just a mapping into another character set, i. e., it represents medical knowledge.

The most important relation used for classification is the hierarchic relation. The following types of classifications are distinguished:

- Monohierarchic classifications are generated by successive refinement and are, typically, one-dimensional. Each concept has exactly one parent.
- Polyhierarchic classifications are generated when the same concept is refined by more than one attribute. Therefore, a concept may have several parents.

Figure 1. Mapping a hierarchic classification onto a numerical code.



Hierarchical relations, in turn, can be divided into:

- Partitive – a total-part relation i. e. "is a relationship".
- Generic – a logical relation of parts i. e. a specialization or a superordinated/subordinated concept.

Monohierarchical classifications are frequently insufficient for complex applications. On the other hand, polyhierarchical classifications tend to be voluminous and redundant and require such an investment for maintenance that, frequently, they become outdated soon after their original design.

In order to realize the advantages of polyhierarchical classification or nomenclature, a significant reduction of the total size without a considerable loss in precision can be reached if there is a semantic model of the domain of discourse. Then, a polyhierarchical classification can be replaced by:

- Several disjoint monohierarchical classifications connected by a semantic model.

If such a model is paralleled by a decomposition of statements into parts belonging to the different dimensions of the model then a polyhierarchical dictionary can be replaced by:

- Several disjoint monohierarchical dictionaries the combination of which is described by a metalanguage.

This concept will be illustrated using SNOMED as an example of such a structure.

SNOMED

SNOMED is a seven-dimensional categorized nomenclature consisting of about 45.000 terms classified into exactly 1 out of 7 semantic dimensions. The German adaptation is about twice as large because multiple synonyms, general classes, and specific morphemes have been added. If modifiers are considered, the number of lexemes rises to more than 200.000. Each dimension is a structure tree with up to five levels generated by considering hierarchic relations. The hierarchy is reflected in the hierarchic code structure. Each code consists of a letter indicating the semantic dimension and of up to 5 characters represented by the digits 0 to 9, A thru F. In Topography, the hierarchy is created by considering the partitive hierarchic relation; in the other dimensions, it is created by considering the generic hierarchic relation.

The SNOMED dimensions are:

Topography	Body parts, body systems, organs and regions. This dimension contains the terms for tissues, organs, body fluids and spaces as well as subcellular structures.
Morphology	Alterations of normal anatomy, whether congenital or acquired after birth, i. e., pathologic anatomy of cells, tissues, organs and systems.
Etiology	Terms used to describe causes of morphological and/or functional alterations such as organisms, chemicals, drugs, physical agents.
Function	Terms for (patho)physiology as well as terms used to describe functional states or processes. Its highest order partitioning into chapters is mainly according to the body systems.
Disease	Disease concepts and syndromes. This is a one-dimensional classification of complex nosological units and is mainly partitioned according to the body systems.
Procedure	Administrative, diagnostic and therapeutic activities available to prevent, relieve or cure diseases. Its highest order partitioning is mainly according to the body systems, i. e., the medical specialties.
Occupation	Classification of occupations borrowed from the International Labor Organization. The dimension is of help in coding environmental and occupational relationships to diseases.

Statement	Topography	Morphology	Etiology	Function	Disease
Tuberculosis of the lung	Lung T28000				Tuberculosis D01880
Lipoma of the trunk	Trunk TY1000	Lipoma M88500			
Pharynx paralysis	Pharynx T60000			Paralysis F80840	
Neurosyphilis	Nervous syst. TX0000				Syphilis D01820
Mumps	Parotid gland T55100				Mumps D04330
Streptococcal pharyngitis	Pharynx T60000	Inflammation M4000	Streptococcus E25000		

Figure 2. SNOMED coding examples

If we do not consider the special dimension **Occupation** the fundamental semantic SNO-MED model is:

A **procedure**
for a **disease** or a **morphologic** alteration
in a **localization**
caused by an **agent**
combined with a **functional** disturbance.

According to this model, a statement is transformed into a **TMEFDP** – or **SNOMED statement** Figure 2. For the analysis it is assumed that each class from one dimension can be combined with each class from another dimension.

It is quite important to note that SNOMED is not a general all purpose classification of medical terms. This statement must not be confused with the fact that the system aims at covering the total information space of medical statements, i. e. SNOMED offers the fundamental information structure which is required for indexing medical language data to a significant degree of detail.

Applications for such a powerful multidimensional nomenclature are found wherever language-based communication takes place. Applications range from clinical documentation and health statistics to knowledge organization and automated translation from one medical language to another.

Data mapped into SNOMED may either be retrieved on the basis of low-level information atoms such as: "fracture of the body of the 12th thoracic vertebra due to fall from a ladder"; or on the basis of high level information aggregates such as: "inflammation of the gastrointestinal tract". To do this, just the hierarchic code structure has to be used.

These examples provide some insight into the attribute structure or deep structure of SNO-MED. From the point of view of indexing, the corresponding problem to solve is the decomposition of each medical concept strictly into their respective SNOMED entities and their position in a SNOMED space (Figure 3).

SNOMED, as it exists, is not an explicit definition of the deep structure but a structured multidimensional list of denotation of the basic entities, i. e., a list of terms and names from which the deep structure has to be derived. Therefore, an indexing procedure must include a set or a sequence of transformations of a given language string which ultimately result in a formulation reflecting the deep structure by using only SNOMED terms in the denotation of the concept. This may be extended by use of the relations between the different dimensions.

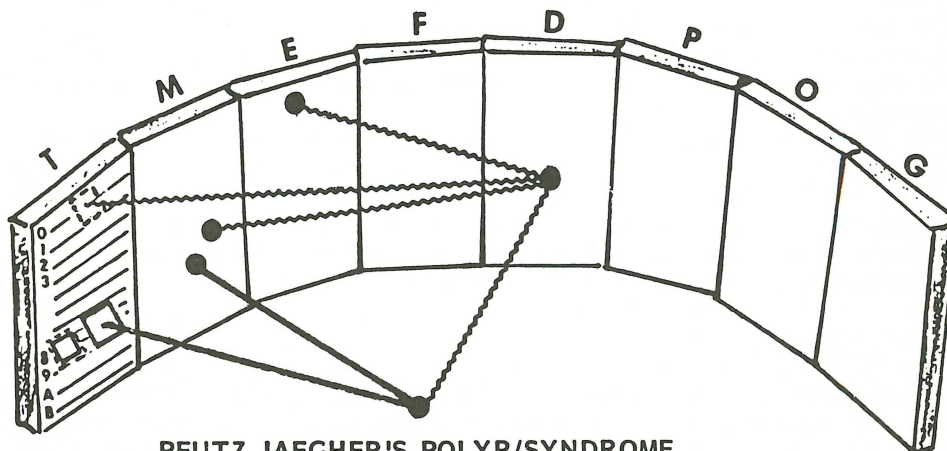
Significant advances in the techniques for automated encoding of medical terms have occurred during the past few years.^{14, 15} Through many have looked at these advances as a panacea

for their information needs, automated encoding has merely served to accentuate rather than solve the general problems inherent in an information storage and retrieval system. In part due to the improved ability to autoencode terms, either automatically or through an interactive technique, the pressure to retrieve information has increased substantially. Correspondingly, the complexity of information requests has increased. This desire for coordinated, detailed data is yet another reason for the creation of a comprehensive medical information system. Although the accuracy of coding will increase with advances in autoencoding techniques, there remains the more fundamental problem of constructing an indexing or coding system so that relationships between terms will be evident and comprehensive retrieval can be achieved.

The principle of multidimensional nomenclatures has been introduced above using SNOMED as an example. Multidimensional nomenclature may open a way out of the dilemma that we need classifications for building intelligent documentation systems and at the same time are confronted with the insufficiencies of these classifications.¹⁶ In order to play the role of an intermediate language sufficiently well, a nomenclature has to meet certain requirements:

- The semantic model must reflect the medical principles of aggregating information, i. e., it must be multidimensional. The selection of dimensions must be such that the structure is elementary enough to serve as a basis for expressing medical information of all kinds and at the same time be simple enough to be handled with the algorithmic instruments that are available now.
- The nomenclature must incorporate as much medical knowledge as possible. For example, it must have the means for the recognition of synonyms, homonyms, and hierarchical relations.

INFORMATION SPACE ----- SNOMED



PEUTZ JAEGER'S POLYP/SYNDROME

T64, M7563

(CR. T67 - Large Bowel;
T01 - skin M57 - pigmentation
T5103 - mouth
E-0103 - autosomal dominant
D 5432 - P.J. Syndrome)

Figure 3.

- The structure of the nomenclature must be free of classificatory elements that lead to a bias which makes mapping into other classifications more difficult or even impossible, e. g. the position of pituitary dwarfism does not prohibit the assignment (classification) of this entity among dwarfism as well as among diseases of the pituitary gland.

These requirements call for the introduction of associative relations, among which cross references across and within dimensions are of most importance. This is absolutely necessary to fully specify and relate information found in a single dimension (monohierarchy) with information that is contained within a given concept in other dimensions (polyhierarchy).

- The nomenclature must be unambiguous. No information unit may be representable by different projections. This rule is crucial.

- The nomenclature must be complete both with respect to the classes and with respect to the terms. In this context it must be kept in mind that a nomenclature is always explicit, i. e., it is a list of terms and not a set of rules, telling where a term has to be positioned.

- The nomenclature must be free of redundancy in order to reduce maintenance problems. In a recent paper, FEINSTEIN¹⁷ outlines four major requirements for a system of taxonomic classification in science. 1) The classification must have a suitable organizing principle. 2) The categories must be labeled with standard titles. 3) The members of each category must be identified with a suitable set of operational criteria. 4) It should be multiaxial. The organizing principle of SNOMED rests on its dimension i. e. where is the lesion (topography); how is the lesion described (morphology); what are its associations or causes (etiology); what functional disturbances does it cause (function); what is the name of the lesion (disease) and what is done about it (procedure). It can be argued that these categories (dimensions) form the natural parts of medical speech providing standard titles for each dimension. Defining operational criteria for inclusion of each element within such a system has as yet not been addressed in a direct way.

Each dimension serves as a data dictionary and provides a road map to a data base. Data elements can be aggregated for the different purposes of acute and long term care, insurance needs, literature retrieval and education. A SNOMED statement houses medical data, this represents static knowledge. Circumstances surrounding the data elements need to be specified and a formal syntax is being developed. A SNOMED statement is an example of an artificial language designed to be efficient, non-ambiguous and non-redundant – one code for one concept.

With advances in computerized natural language processing there will be a need for coordination among data bases, expert systems and the output from these language processors. A common taxonomy will be a prerequisite for interaction among the three.¹⁸ This search for common features for representing knowledge becomes even more important when different kinds of systems are interconnected. This commonality is needed to support queries or commands and to format the data used to construct such systems.

Automated Indexing

There is a substantial agreement among knowledgeable users that the SNOMED model outlined above may serve as the structural basis for housing information generated by natural language processors. Natural language processors must perform several functions to transform language into this structured form retaining the essential meaning of the statement in the process. These include a morphologic analysis of each word or lexical unit, a syntactic analysis of each unit to determine the relationship of each unit within the statement and as a result of the above two activities determine its semantic content. This result can be expressed as a discrete point in a SNOMED space where each of the projected dimensions are linked by appropriate qualifying and linkage terms i. e. its syntax and meaning. Professor WINGERT has virtually completed the first of these functions namely morphologic analysis of

each lexical unit found in a medical statement. These lexical units were matched to the underlying SNOMED lexicon and the proper codes assigned to each dimension of a SNOMED space. To achieve this a systematic and comprehensive elaboration of words and word parts are required i. e. a morpheme and suffix family dictionaries together with a set of rules that determine all possible constructions for each lexical unit. Professor WINGERT virtually single-handedly completed this work and it stands as a monument to him. His programs break up a diagnostic statement of restricted lengths (137 characters) into an unordered set of words and word parts; the rule dictionary then selects all possible choices of words from these fragments and presents to the user those most likely to match the underlying SNOMED lexicon. This procedure works with great efficiency and nearly always (greater than 95 %) finds a correct match. Frequently however multiple and ambiguous matches are presented for selection.

This approach works well for short and well-standardized statements. Part of its attraction, however, is based upon what it can do with considerably longer and more complex statements. In these, it has been observed that the correspondence between the noun phrase syntax and the formulation of dictionary entries is characteristically very poor. Medical syntax is most prolific in paraphrastic variation. Discontinuous sequences, in which word/segments of a dictionary entry may be found in syntactically unrelated parts of a diagnostic utterance are commonplace.¹⁹

To improve this basic design, one adds imperfect heuristics for weeding superfluous matches based upon the principle of longest match. Semantic rules can be expanded before dictionary matching by doing a bit of immediate constituent syntactic analysis to function with greater precision after matching, "semantic rules" can be expanded to rule out specific matches as redundant or semantically incompatible combinations. Here, one is essentially fine tuning the unpredictable results of the principle of longest match. By such means, one may with considerable effort, arrive at a system which works well at a higher degree of accuracy, without having done any serious syntactic analysis of the diagnostic utterance.

Syntactic problems include identifying the metastatic and primary sites of a histologically typed malignant tumor, and determining whether adjectives and some nouns are properly or improperly matched with words from a neighboring noun phrase. In these typical cases, the reasoning which leads to satisfactory decisions requires both syntactic information from the utterance, semantic information from lexical lookup and semantic rules. Syntactic analysis can also be useful prior to dictionary lookup by selecting smaller subsets of word/segments of the utterance for which syntactic analysis can reasonably predict co-membership in word/fragment sets of dictionary entries.²⁰ It seems reasonable to learn to cope well with these relatively simple sorts of syntactic problems in view of our ambition to eventually deal with more complex, syntactically rich medical documents such as clinical reports.

Conclusion

These then are an outline of the tasks that Professor WINGERT addressed and to which he devoted so much of this time, energy and abundant talent. His goals are perhaps best summarized in the following quote from the late DR. MARSDEN BLOIS in an essay written shortly before his death last year.

The medical practitioner needs to be able to harness the tools of reasoning better in order to apply them to a mixture of low-, middle-, and high-level data. This is essential if physicians are to range back and forth, consciously and effectively, from the mathematical descriptions of atomic and molecular events to the statistical associations exhibited by complex biologic systems, and to the natural-language descriptions at the clinical and behavioral levels. But the first step will be to recognize the existence of this structure and appreciate the multiple levels of diagnosis and treatment - some of which will always be more important than others, but none of which can be left unattended. To succeed fully in this effort will require a better

understanding of the relations between computation and language and between formal logic and common sense. No small tasks! Yet it is just these issues that are central to modern information science and its research program, particularly as it attempts to reach a practical understanding of the structure of medical knowledge so that we may better describe and reason in order to preserve health or to cure.²¹
We must continue this work.

References

1. MAYR, E.: Biological Classification: Toward a Synthesis of Opposing Methodologies. *Science* 1981; 214 : 510-516.
2. CÔTÉ, R. A. and ROBBOY, S.: Progress in Medical Information Management. *JAMA* 1980; 243 : 756-762.
3. CÔTÉ, R. A.: International Classifications for Health and Disease: The Expandable Common Core Concept. *Medical Informatics* 1983; 8 : 5-16.
4. International Classification of Diseases World Health Organization. Geneva, 1978.
5. THOMPSON, B. J.: Implications of an Interactive Encoding System for ICD-9-CM. Eleventh Annual Conference of the Society for Computer Medicine, Washington, DC, 1981, 11 pp.
6. Standard Nomenclature of Diseases and Operations. McGraw Hill, New York, 1959.
7. HUTCHINS, W. J.: Languages of Indexing and Classification. Peter Berekgrinus, Ltd, Southgate House, Stevenage, Herts., England, 1975.
8. SPARCK JONES, K. and KAY, M.: Linguistics and Information Science. Academic Press, New York, 1983.
9. Medical Subject Headings. Index Medicus 1983; 24 : 1-877.
10. SOERGEL, D.: Indexing Languages and Thesauri: Construction and Maintenance. Melville Publishing Company, Los Angeles, 1974.
11. Systematized Nomenclature of Medicine. 2nd ed., College of American Pathologists, Chicago, 1979.
12. WINGERT, F.: Medical Linguistics: Automated Indexing into SNOMED. *Critical Reviews in Medical Informatics*. 1988; 1: 333-403.
13. WINGERT, F., ROTHWELL, D., CÔTÉ, R.: Automated Indexing into SNOMED and ICD. IMIA - WG 6. Role of Informatics in the Classification and Coding of Health Data. Geneva, Switzerland. September 1988.
14. SAGER, N., FRIEDMAN, C., LYMAN M.: Medical Language Processing. Addison-Wesley Pub. Co. Reading, MA 1987.
15. ROTHWELL, D. J., HAUSE, L. L.: SNOMED and Microcomputers in Anatomic Pathology. *Med. Inform.* 1983; 8 : 23-31.
16. PACAK, M. G., DUNHAM G. S.: Computer & Medical Language. *Med. Inform.* 1979; 4: 13 - 27.
17. FEINSTEIN, A.: ICD, PDR and DRG. *Arch. Int. Med.* 1988; 148 : 2269.
18. SOWA, J.: Working Group on Data bases. (WG 26) Elsevier, North Holland 1989.
19. DUNHAM, J.: Personal Communication.
20. DUNHAM, G. S., HENSON, D. E., PACAK, M. G.: Three Solutions to Problems of Categorized Medical Nomenclature. *Meth. Inform. Med.* 1984; 23: 87 - 95.
21. BLOIS, M. S.: Medicine and the Nature of Vertical Reasoning. *NEJM* 1988; 318 : 847-851.

Anschrift des Verfassers: David J. Rothwell, M. D., Clinical Professor of Pathology, The Medical College of Wisconsin and Columbia Hospital, 2025 East Newport Ave., Milwaukee, Wisconsin 53 211, USA.

SNOMED-Anwendungen

J. Berger, E. Hultsch, R. Thurmayr

Zusammenfassung

Das von FRIEDRICH WINGERT entwickelte Programm generiert zu einer medizinischen Aussage automatisch die zugehörigen SNOMED-Indizes und bildet diese in die ICD-9 ab. Es werden die Anwendungen dieses Programms beschrieben und über die vorliegenden Erfahrungen berichtet. Insbesondere wird das Programm in den Universitätskliniken Eppendorf seit 1986 in der Dokumentation (SNOMED) und der Diagnosenstatistik (ICD-9) erfolgreich eingesetzt. Die Arbeiten mit dem Programm und an SNOMED ergänzen einander. Zur Fortführung der Arbeiten von FRIEDRICH WINGERT ist es einerseits notwendig, in Zusammenarbeit mit „The College of American Pathologists“ SNOMED weiterzuentwickeln. Andererseits muß eine Organisationsform gefunden werden, so daß Pflege und Wartung des Programmpakets gewährleistet sind.

Abstract

FRIEDRICH WINGERT developed a program for the automatic generation of SNOMED-indices and the transformation into ICD-9. The applications of this program and the acquired experiences are related. Since 1986 the program has been employed in the documentation (SNOMED) and statistics of diagnoses (ICD-9) especially in the Universitätskliniken Eppendorf with great success. Working with the program and development of SNOMED complete each other. In order to continue the work of FRIEDRICH WINGERT it is necessary to develop SNOMED in cooperation with "The College of American Pathologists". On the other side there has to be found a form of organisation so that management and maintenance of the program set are guaranteed.

Key words

SNOMED-applications, Automated Indexing System

Die manuelle Zuordnung von Kodes zu medizinischen Texten ist bei der Systematischen Nomenklatur der MEDizin (SNOMED) nicht viel aufwendiger als bei differenzierteren Klassifikationen. Einer Anwendung steht in beiden Fällen oft entgegen, daß man zur Kodierung gut ausgebildetes (und teures) Personal benötigt. R. P. Baumann unterstützt die Kodierung von Biopsie- und Autopsiebefunden durch ein einfaches Dialogprogramm mit The-saurusfunktionen.

Das von WINGERT entwickelte Programm zur Indexierung analysiert im Klartext formulierte medizinische Phrasen nach morphologischen Gesichtspunkten [5] und ordnet ihnen SNOMED-Indizes zu. Bei Diagnosebezeichnungen werden die SNOMED-Indizes in die

(3-stelligen) Codes der ICD-9-Klassifikation abgebildet. Zudem werden die Codes für die Histologie von Tumoren nach der ICD-0 bereitgestellt, da die SNOMED-Indizes im 8. und 9. Kapitel der Morphologie-Dimension mit der ICD-0-Klassifikation identisch sind.

Bei Beachtung einiger einfacher Regeln liefert das Programm zu eingegebenen medizinischen Klartexten in über 90 % aller Fälle die richtigen SNOMED-Indizes, so daß die Inhalte der Texte für medizinische Dokumentationen und als Basis für wissenschaftliche Auswertungen zur Verfügung stehen.

1. SNOMED-Indizes und ihre Generierung

Der erste Ansatz zu einer Nomenklatur klinischer Diagnosen wurde von dem „College of American Pathologists“ mit der „Systematized Nomenclature of Pathology (SNOP)“ unternommen. SNOP ist eine 4-dimensionale und in jeder Dimension hierarchisch strukturierte Nomenklatur. Die breite Akzeptanz führte zur Übertragung des Ansatzes auf den gesamten medizinischen Bereich: SNOMED unterscheidet sich von SNOP durch einen höheren Differenzierungsgrad und die Einführung weiterer Dimensionen, die für einen Einsatz in der klinischen Medizin notwendig waren (SNOMED I) [1].

WINGERT hat diese Nomenklatur übersetzt, wesentlich ergänzt (die Anzahl der aufgeführten Termini ist etwa doppelt so groß) und wichtige, für eine automatische Indexierung notwendige Erweiterungen durchgeführt (SNOMED II) [3].

Die SNOMED spaltet jede medizinische Aussage in bis zu 7 „Dimensionen“ auf:

- Topographie (etwa 18 000 Einträge),
- Morphologie (etwa 15 000 Einträge),
- Ätiologie (etwa 18 000 Einträge),
- Funktion (etwa 13 000 Einträge),
- Krankheit (etwa 16 000 Einträge),
- Prozedur (etwa 7 000 Einträge) und
- Beruf (etwa 1 400 Einträge).

Das den Dimensionen von SNOMED zugrundeliegende Aussagemodell lautet: Prozedur (P) wegen morphologischer Veränderung (M) mit Funktionsstörung (F), bedingt durch ein ätiologisches Agens (E) an einer Lokalisation (T) [verursacht durch Ausübung eines Berufs (J)].

Beispiel: Die Aussage „Oesophagusbougieung wegen Verätzung mit Salzsäure“ wird durch das Aussagemodell abgebildet auf „Bougierung (P) wegen Verätzung (M), bedingt durch Salzsäure (E) am Oesophagus (T)“.

Die Krankheit (D) ist keine echte Dimension, sondern kann durch eine Aussage des Aussagemodells ersetzt werden.

Beispiel: „Tuberkulose“ entspricht „Granulom (M) mit Fieber (F), bedingt durch Mycobacterium tuberculosis (E) an der Lunge (T)“.

Jede der SNOMED-Dimensionen ist in den Begriffen hierarchisch (mit bis zu 5 Ebenen) strukturiert (Relation: Oberbegriff – Unterbegriff). Daher entspricht jedem Begriff ein maximal 5-stelliger Kode.

Beispiel: Der Aussage „Bougierung (P) wegen Verätzung (M), bedingt durch Salzsäure (E) am Oesophagus (T)“ entspricht der Kode: (T 62000, M 11150, E 50220, F, P 17230, [JJ]).

Zu SNOMED gibt es zumindest keine prinzipielle Alternative:

- Jede medizinische Nomenklatur muß mehrdimensional sein, da sonst die Anzahl der Termini zu groß wird und damit Wartung und Pflege praktisch unmöglich sind.
- Das Aussagemodell von SNOMED entspricht dem semantischen Aufbau medizinischer Information. SNOMED kann daher weitestgehend die Prinzipien der Bildung medizinischer Termini ausnutzen.

- Der durch den hierarchischen 5-stelligen Kode zur Verfügung stehende Spezifizierungsgrad ist nach bisheriger Erfahrung für klinische Anwendungen voll ausreichend. Dadurch ist gewährleistet, daß alle klinisch relevanten Klassifikationen (wie etwa ICD) als Vergrößerungen dargestellt werden können.

Die Algorithmen zur automatischen Generierung der SNOMED-Indizes beruhen auf einer morphologischen Analyse des einzugebenden Klartextes. Die Algorithmen sind in einer ganzen Reihe von Arbeiten WINGERTS beschrieben (etwa [4, 5, 6]). Im Gegensatz zu diesen Veröffentlichungen wird in dem Programm für die Segmentierung das folgende Wortmodell benutzt:

$$w = r_1 s_1 \dots r_n - 1 s_n - 1 r_n t \quad \text{mit } n \geq 1 \text{ und}$$

- r_i : Wortstämme,
- s_i : Suffixe (Derivations-, Fugenmorpheme) und
- t : Flexionssuffix (Derivations-, Flexionsmorpheme).

Es hat sich gezeigt, daß bei der Analyse von medizinischen Phrasen dieses vereinfachte Wortmodell bei geringerem Pflegeaufwand zu kürzeren Verarbeitungszeiten führt.

2. Abbildung von SNOMED-Aussagen in Klassifikationen

Die Kodierung von Diagnosen nach SNOMED ist sicherlich dann die sinnvollste Alternative, wenn formale Abbildungsstrukturen benötigt werden, die möglichst viel der Ursprungsinformation beinhalten. Dies gilt i. a. dann, wenn Diagnosen für eine Dokumentation kodiert und die Retrieval-Möglichkeiten nicht eingeschränkt werden sollen.

Bei Anwendungen, bei denen der Zweck vorher festliegt, benötigt man meist Klassifikationen. Die Effizienz einer Klassifikation ist abhängig von der Anwendung, so etwa, ob die Dokumentation für

statistische Zwecke (z. B. „wie oft wurden im letzten Jahr Frakturen der unteren Extremität behandelt?“),

Forschung (z. B. Vergleich zweier Therapieformen bezüglich Nebenwirkungen) oder

die Analyse der Ursachen von Krankheitskosten

gedacht ist. Keine der bekannten Klassifikationen ist für alle Zwecke geeignet. Neben den international anerkannten Klassifikationen gibt es daher eine große Anzahl national, zumeist aber nur regional gültiger Klassifikationen, die über den jeweiligen Anwendungsbereich hinaus keine größere Verbreitung gefunden haben.

Die Effizienz einer Klassifikation ist nicht nur vom Anwendungsgebiet, sondern auch von der Konsistenz der Zuordnung von medizinischen Aussagen zu Codes abhängig. Neben einer EDV-Unterstützung ist dies nur möglich, wenn die Pflege und Wartung einer Klassifikation durch Fachleute zentral durchgeführt und das die Klassifikation anwendende Personal laufend geschult wird. In den letzten Jahrzehnten ist es aus Gründen der nationalen und internationalen Vergleichbarkeit immer wichtiger geworden, für Verschlüsselungen international anerkannte und verbreitete Klassifikationen zu benutzen.

Wie jede Klassifikation ist die ICD-9 eine Vergrößerung von SNOMED: Für die meisten Diagnosen bedeutet der zugehörige ICD-9-Kode einen Informationsverlust gegenüber der zur Diagnose gehörenden SNOMED-Aussage, d. h. die Eingangsphrasen und die SNOMED-Aussagen erlauben eine stärkere Spezifikation als der ICD-9-Kode.

Zu jedem SNOMED-Eintrag gehört ein ICD-Eintrag und zu jedem ICD-Eintrag eine Liste von SNOMED-Einträgen („region“), nur in seltenen Fällen findet man eine 1:1-Zuordnung. Diese Abbildung von SNOMED auf ICD kann mit Hilfe von Transformationstabellen durchgeführt werden, da SNOMED eine systematisierte Nomenklatur ist und damit eine solche Formalisierung unterstützt.

Bei der Anwendung der ICD-9 ist eine Reihe von Regeln zu beachten:

- Stimmen zwei einer Diagnose zuzuordnenden ICD-9-Kodes in den ersten 3 Stellen überein, so hat der ICD-9-Kode mit den Ziffern 0 bis 7 den Vorzug vor 8 und dieser wieder vor 9. Diese Regel folgt aus der meist eingehaltenen Struktur des ICD-9, nach der die Subkategorie .8 für „Sonstiges“ und die Subkategorie .9 für „ohne nähere Angaben“ verwendet werden.
- Stimmen zwei einer Diagnose zuzuordnenden ICD-9-Kodes in den ersten 3 Stellen nicht überein, dann ist auch hier der Kode mit dem kleineren Wert zu wählen, da in den Kategorien ebenfalls das Prinzip gilt, daß „Sonstiges“ und „ohne nähere Angaben“ erst nach Aufzählung der explizit genannten Ausprägungen der jeweiligen Kategorie aufgeführt werden.

Wegen dieser Mehrdeutigkeiten eignen sich als Hilfe bei der Definition der „regions“ von SNOMED besonders die Phrasen des alphabetischen Teils, da hier auch Diagnosen der Restmengen explizit aufgeführt sind: Alle im alphabetischen Teil aufgeführten Diagnosen werden nach SNOMED indexiert und nach den einzelnen ICD-9-Kodes geordnet.

Eine weitere Schwierigkeit besteht darin, daß die ICD-9 in den ersten 3 Kapiteln ätiologisch und in den nachfolgenden organbezogen aufgebaut ist. Daher kann einer Diagnose oftmals sowohl ein Kode zugeordnet werden, wenn man sich auf die Ätiologie der Erkrankung bezieht, als auch ein anderer Kode, wenn das erkrankte Organ Vorrang hat. Solche Stellen sind in der ICD-9 mit „+“ (Kreuz) und „*“ (Stern) gekennzeichnet. Auch in diesen Fällen erhält der Kode mit dem kleineren numerischen Wert den Vorzug – wenn man der Empfehlung der WHO oder der Vorschrift der Bundespflegesatzverordnung folgt – daß „Ätiologie vor Organ rangiert“.

Beispiel: Der Diagnose „Blasentuberkulose“ wird in der ICD-9 der +-Kode

016.1 Tuberkulose sonstiger Harnorgane

und der *-Kode

595.4 Zystitis anderswo klassifizierter Krankheiten

zugeordnet. Nach Empfehlung der WHO oder der Bundespflegesatzverordnung ist der Kode „016.1“ zu wählen.

Im Gegensatz zu den Empfehlungen der WHO erscheint es aber für bestimmte Anwendungen durchaus sinnvoll, dem *-Kode den Vorzug zu geben. Es dürfte am sinnvollsten sein, beide Möglichkeiten vorzusehen und dem Anwender die Entscheidung zu überlassen, für welche der Möglichkeiten er sich entscheidet.

Die diesen Transformationstabellen zugrundeliegenden Algorithmen wurden von WINGERT veröffentlicht [7]. Das Programm bietet damit die Möglichkeit, einen medizinischen Klartext einzugeben und automatisch die zugehörigen SNOMED-Indizes als Basis für Dokumentation und wissenschaftliche Auswertungen und den zugehörigen ICD-Kode für statistische und/oder administrative Zwecke zu erhalten.

3. Regeln bei der Klartexterfassung

Bei der Texteingabe sind einige einfache Regeln zu beachten. So muß dem zu indexierenden Text als Kennbuchstabe entweder „D“ für Diagnose oder „O“ für therapeutisches, diagnostisches oder administratives Handeln vorangestellt werden. Der dann unmittelbar an den Kennbuchstaben anschließende Text darf zwischen zwei Worten nur ein Leerzeichen aufweisen und maximal 80 Zeichen lang sein. Da bei der Textaufbereitung im Rechner vor der Indexierung nicht-sinntragende Worte, wie beispielsweise Artikel, eliminiert und Endungen abgeschnitten werden, können diese bei der Texteingabe weggelassen werden. Da Hilfsverben in Befunden keine Information enthalten, können sie ebenfalls bei der Texteingabe weggelassen werden.

Unterschiedliche Schreibweisen eines Wortes wie z. B. Karzinom mit K und z, Carzinom mit C und z, Carcinoma mit C und c oder Mikro mit k, Micro mit c oder Cornea mit C, Kor-

nea mit K oder synonyme Endungen wie om/oma, ose/osis, ie/ia, iase/iasis stören den Indexierungsvorgang nicht. Es ist aber ansonsten darauf zu achten, daß die Eingangssphrasen der im Programm geforderten Schreibweise (Groß- und Kleinschreibung, Umlaute, scharfes ß und griechische Buchstaben) entsprechen.

Bei der Verwendung von Vorzugsbezeichnungen müssen alle Phrasen entfernt werden, in denen „Sonstiges“ oder „ohne nähere Angaben“ stehen, da beide lediglich auf eine fehlende Spezifikation oder auf eine Restmenge hinweisen.

Informationsqualifikatoren wie „Zustand nach“ oder „Verdacht auf“ sollten als Abkürzungen „Z. n.“ bzw. „V. a.“ dem Text vorangestellt werden. Für gängige diagnostische und therapeutische Bezeichnungen sind auch die allgemein üblichen Abkürzungen in den Lexika enthalten, so daß es empfehlenswert ist, diese bei der Texteingabe zu benutzen, um sowohl Zeit zu sparen als auch Rechtschreibfehler zu vermeiden.

Enthält ein Satz mehrere diagnostische oder therapeutische Angaben, so empfiehlt es sich, um Eindeutigkeiten bei der Indexierung zu erhalten, diese jeweils auch als getrennte Texte indexieren zu lassen.

Ist eine Indexierung wegen eines auftretenden Segmentierungsfehlers nicht möglich, so meldet das Programm die Nummer des Wortes, in dem der Fehler auftrat. Es sollte dann als erstes dieses Wort auf Rechtschreibfehler überprüft werden, da Schreibfehler bei der Texteingabe meist der Grund für das Auftreten des Fehlers sind.

Fällt die Indexierung bei der morphometrischen Analyse nicht eindeutig aus, so werden alle möglichen Indexierungsvorschläge durchnummeriert ausgegeben. Dies kann dann der Fall sein, wenn aufgrund des Algorithmus mehrere Zuordnungen der Wortsegmente syntaktisch möglich sind.

Beispiel: Für die Diagnose „Adeno-Ca Magen“ erhält man die beiden Indexierungsvorschläge:

- 1) T63000 Ventriculus
M81403 Adenocarcinoma
I151 bösartige Neubildung des Magens,
- 2) T63040 G11 gastricae propriae
M80103 Karzinom
I151 bösartige Neubildung des Magens.

Der erste Indexierungsvorschlag ist der semantisch richtige, der zweite Vorschlag beruht darauf, daß bei der morphometrischen Analyse der Wortstamm Adeno in Verbindung mit dem Magen interpretierbar ist.

Mehrere Indexierungsvorschläge können auch dann auftreten, wenn der dem Programm angebotene Text nicht eindeutig formuliert ist.

Beispiel: Bei der Diagnose „Ulnarisdurchtrennung“ kann sowohl der Nervus ulnaris als auch die Arteria ulnaris betroffen sein. Das Programm beendet den Indexierungsprozeß mit den beiden Vorschlägen:

- 1) TX9170 N. ulnaris
M14300 Vulnus sectum
I955 Verletzungen der peripheren Nerven des Schultergürtels und der oberen Extremitäten
I869 Innere Verletzung von n. n. bez. oder mangelhaft bez. Organen
- 2) T47200 A. ulnaris
M14300 Vulnus sectum
I869 Innere Verletzung von n. n. bez. oder mangelhaft bez. Organen

Beachtet man bei der Texteingabe diese einfachen Regeln, so können 90 bis 95 % aller medizinischen Aussagen eindeutig indexiert werden.

Die häufigsten Fehler, die beim Indexieren natürlichsprachlicher medizinischer Aussagen auftreten, sind

- fehlende und inkonsistente SNOMED-Einträge,
- Unzulänglichkeiten im Verarbeitungsprozeß,
- unzureichende morphologische Segmentierung,
- Erzeugung von überflüssigen Mehrdeutigkeiten durch:
 - fehlende syntaktische Informationen bei der Analyse,
 - fehlende semantische Informationen bei der Analyse,
 - unbeschränkte rekursive Einsetzung von Indexeinträgen,
- mangelnde Wissensquellen wie
 - medizinische Wörterbücher und Lexika,
 - pragmatisches und situationsgebundenes Wissen.

4. Einsatz in der Routine

Das von WINGERT entwickelte Programmsystem zur Indexierung von Medizinischen Aussagen nach SNOMED und Transformation in die (3-stellige) ICD-9 hat sich in den Universitätskliniken Eppendorf seit 1986 im Routineeinsatz gut bewährt. In dieser Zeit wurden ca. 80 000 Texte indexiert.

Die den medizinischen Aussagen zugeordneten SNOMED-Kodes werden im Rahmen der Dokumentation verwendet. Durch die SNOMED-Kodes können die Texte sehr viel detaillierter beschrieben werden, als dies mit einer Kodierung nach einem einachsigen Schlüssel, wie ihn beispielsweise ICD-9 darstellt, möglich ist. Dies bietet nicht zu vernachlässigende Vorteile für wissenschaftliche Auswertungen der Daten. Die ICD-9-Kodes werden zusätzlich im Rahmen der Diagnosenstatistik entsprechend § 16 der Bundespflegesatzverordnung benötigt.

4.1. Datenfluß von der Station zur Dokumentation

Seit Einführung der Diagnosenstatistik erfolgt die Dokumentation an den Universitätskliniken Eppendorf zentral rechnergestützt nach SNOMED unter dem Einsatz der von WINGERT konzipierten und programmierten Indexierungsalgorithmen. Der Datenfluß ist in Abb. 1 dargestellt. In der zentralen Patientenaufnahme werden die Stammdaten aller im Klinikum stationär aufgenommenen Patienten an Bildschirmarbeitsplätzen erfaßt. Im Rechenzentrum des Klinikums werden neben den anderen für die Station bestimmten Unterlagen zwei Dokumentationsetiketten erstellt, auf denen neben der Stationsnummer die Initialen des Patienten und die eindeutige Kennziffer (KG-Nummer) für den momentanen stationären Aufenthalt des Patienten abgedruckt sind. Auf jeder Station befinden sich Blankoformulare für die Basisdokumentation. Bei jeder Entlassung bzw. Verlegung des Patienten aus der Fachabteilung ist der zuständige behandelnde Arzt verpflichtet, die für die Erstellung der Diagnosenstatistik benötigten Angaben auf einem Dokumentationsbogen einzutragen. Für jeden Patienten können maximal sieben diagnostische und vier therapeutische Angaben auf das Blatt geschrieben werden. Durch Aufkleben des der Station zugeordneten Dokumentationsetikettes auf dieses Formular wird der Schreibaufwand zur Patien-

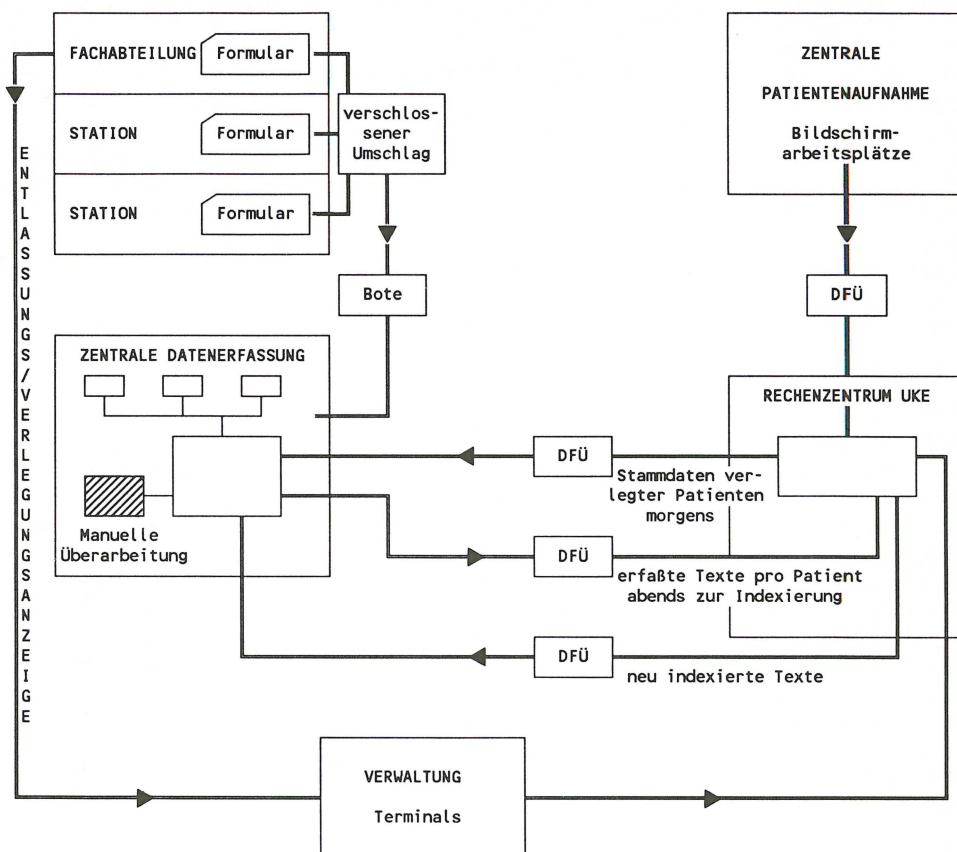


Abbildung 1: Datenfluß im Rahmen der Diagnosenstatistik in den Universitätskliniken Eppendorf (UKE)

tenidentifikation minimiert, und Schreibfehler bei der Übertragung der Identifikationsziffer werden vermieden. Durch die Anonymisierung wird dem Datenschutz Rechnung getragen.

Die ausgefüllten Dokumentationsbögen werden in verschlossenen Umschlägen täglich der zentralen Datenerfassung überbracht. In dieser Einheit sind drei Hilfskräfte (Arzthelferinnen) und ein approbierter Mediziner tätig. Die Erfassung der Klartexte erfolgt von den Hilfskräften an Bildschirmarbeitsplätzen unter Einsatz des mehrplatzfähigen Siemens-Rechners MX2.

Die Stammdaten aller zur Dokumentation anstehenden Patienten werden jeweils morgens vom Zentralrechner abgerufen und in einer Datenbank (INFORMIX) auf dem MX-System abgespeichert. Durch Eingabe der auf dem Etikett abgedruckten KG- und Stationsnummer wird der entsprechende Stammsatz des Patienten für die Dokumentation aktiviert, und auf dem Bildschirm erscheint eine Erfassungsmaske. Die Datenerfasserinnen sind gehalten, nur Texte einzugeben, deren Sinngehalt sie auch verstehen. Im Zweifelsfalle steht ihnen für Rückfragen der die Dokumentation vor Ort betreuende Arzt zur Verfügung, der gegebenenfalls auch Rücksprache mit dem für die Dokumentation zuständigen Arzt der Klinik hält.

4.2. Verarbeitung der Texte

Alle an einem Tag erfaßten Texte (jeweils 1 Datensatz pro Patient) werden zusammen mit der jeweiligen KG-Nummer des Patienten abends zum Zentralrechner des Klinikums übertragen und in einer dafür bestimmten Datei (PDO-Datei) abgelegt.

Im Zentralrechner existiert eine Thesaurus-Datei (TST-Datei), die alle bisher aufgetretenen Diagnose- und Therapietexte zusammen mit den zugehörigen SNOMED-Indizes, ICD-9-Kodes (nur für Diagnosen) und einer identifizierenden Thesaurus-Nummer enthält.

Die Indexierung der Texte wird während der Nacht als Batch-Lauf durchgeführt:

- Wird ein neu erfaßter Text in der TST-Datei gefunden, so wird die Thesaurus-Nummer in die Stammdatendatei übertragen und der Text in der PDO-Datei gelöscht.
- Wird ein neu erfaßter Text nicht in der TST-Datei gefunden, so wird der Text dem Indexierungsprogramm übergeben. Ein- und Ausgabe werden in einer Datei gespeichert und zur Freigabe an die zentrale Datenerfassung übermittelt.

Die Hauptaufgabe des in der Dokumentation tätigen Mediziners besteht darin, die von den Algorithmen gelieferten Indexierungen auf Richtigkeit und Vollständigkeit zu überprüfen und ggf. zu korrigieren. Nach der Freigabe der Indexierung eines neuen Textes wird diesem eine neue Thesaurusnummer zugeordnet und der Text zusammen mit den SNOMED-Indizes und dem evtl. ICD-9-Kode in die TST-Datei aufgenommen. Die neue Thesaurus-Nummer wird in die Stammdatendatei übertragen und der Text in der PDO-Datei gelöscht.

Durch die Zwischenschaltung der Thesaurus-Datei in den Indexierungsprozeß wird eine erhebliche Arbeitserleichterung erreicht, da einmal als richtig indexiert eingestufte Texte bei ihrem erneuten Auftreten nicht wieder indexiert und auf die Richtigkeit überprüft werden müssen. Durch die gewählte Dateistruktur ist darüber hinaus auch gewährleistet, daß nach einer möglichen Revision der Nomenklatur eine schnelle Reindexierung stattfinden kann, da nur die in der TST-Datei abgelegten Texte neu indexiert werden müssen und Änderungen in der Stammdatendatei nicht notwendig sind. Zusätzlich zu den bisher aufgeführten Dateien gibt es für bestimmte Anwendungen eine TYS-Datei, in der zu allen SNOMED-Indizes die zugehörigen Texte abgespeichert sind.

5. Andere Anwendungen

Von den durch das Programm erzeugten SNOMED-Indizes ausgehend, gibt es außerhalb der Dokumentation eine ganze Reihe von Anwendungen:

- Untersuchungen über Tiefe und Konsistenz einer bestehenden Klassifikation oder bei der Generierung neuer Klassifikationen oder
- Analyse der Überschneidungen zweier Klassifikationen und Entwicklung von Umsetzungstabellen als Hilfen für die manuelle Umkodierung.

5.1. Klassifikation und deren Konsistenz

Die von WINGERT entwickelte algorithmische Lösung bietet vielfache Vorteile. Dies ist darin begründet, daß die Menge der SNOMED-Aussagen als „medizinische Sprache“ unter Verwendung von normierten (und dadurch eindeutig definierten) Termen aufgefaßt werden kann. Jede Klassifikation ist dann ähnlich wie die ICD-9 eine Vergrößerung des durch SNOMED-Aussagen definierten semantischen Raums.

In Abschnitt 2 wurde bereits auf die Nutzung des alphabetischen Teils der ICD-9 zur Erzeugung einer Transformationstabelle von SNOMED in die ICD-9 hingewiesen. Innerhalb jeder Dimension sind die SNOMED-Indizes weitgehend hierarchisch aufgebaut, d. h. je weniger Endnullen ein Index besitzt, um so spezifischer ist er. Die Häufigkeit der SNO-

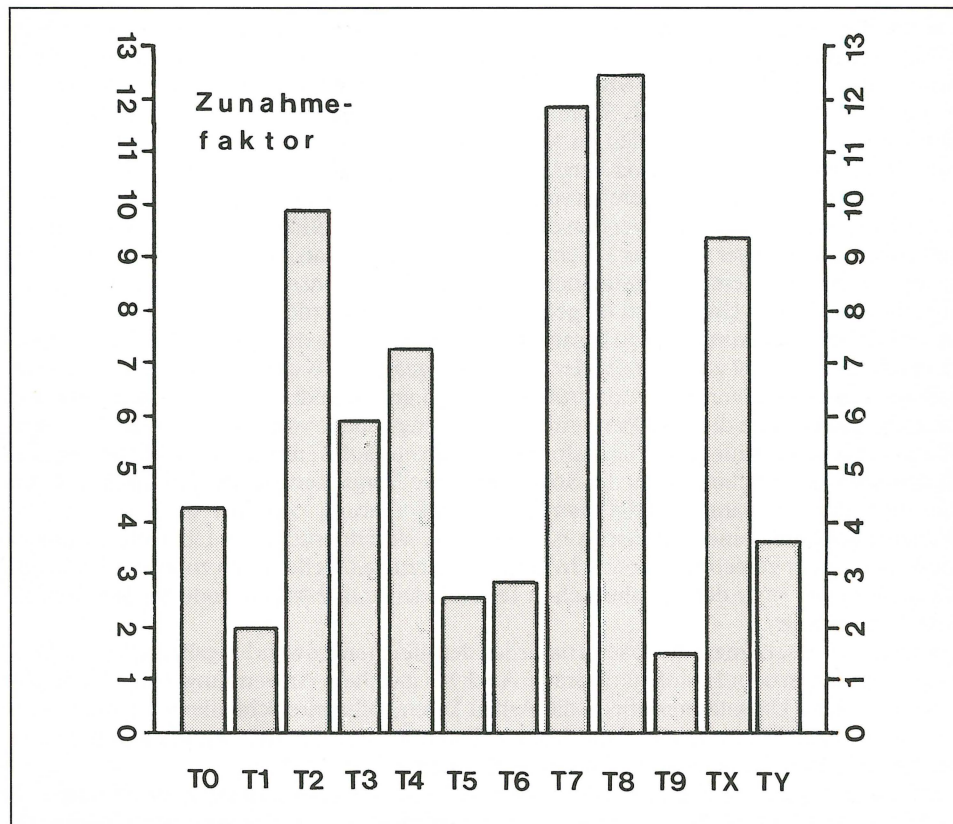
MED-Indizes für jede der SNOMED-Dimensionen auf den verschiedenen Hierarchiestufen ermöglicht die Analyse der Tiefe des Schlüssels. So kann man etwa untersuchen, wie viele SNOMED-Indizes innerhalb eines Organsystems oder eines Organs in der geprüften Klassifikation vorhanden sind. Ein solches Vorgehen, das sich auf alle für die Klassifikation wichtigen SNOMED-Dimensionen erstreckt, deckt sehr rasch fehlende Tiefen eines Schlüssels auf (s. Abb. 2).

Eine gewisse Schwierigkeit bereitet dabei die Tatsache, daß in SNOMED auch faktorielle und mehrfach faktorielle Strukturen auf einen hierarchischen Code abgebildet sind und im Ausdruck nicht besonders gekennzeichnet werden.

Existierende Klassifikationen wie etwa die ICD-9 sind historisch gewachsen, zu einem Teil nur unter Berücksichtigung dieses Aspekts interpretierbar, und in den einzelnen Kategorien sind oftmals Ausnahmen aufgeführt. Diese mangelnde Konsistenz einer Klassifikation führt immer auch zu einer mangelhaften Konsistenz bei der manuellen Zuordnung von Codes.

Abbildung 2: Zunahmefaktoren für topographische Angaben (SNOMED) im allgemeinen Teil des Operationsschlüssels nach SCHEIBE von Version 1 auf Version 2.

TO: Haut und Hautanhangsorgane, T1: Bewegungsapparat, T2: Atmungsorgane, T3: Herz, T4: Gefäßsystem, T5: Verdauungstrakt, T6: Magendarmtrakt, T7: Harnsystem, T8: weibl. Geschlechtsorgane und Fetus, T9: Endokrinium, TX: Nervensystem, TY: topographische Regionen



Schlüsse auf die Konsistenz der ICD-9 können einerseits aus den nach SNOMED-Indizes geordneten Diagnosen des alphabetischen Teils und andererseits aus den Transformationsregeln für die Umsetzung der SNOMED-Indizes in den ICD-9-Kode gezogen werden: Fehlen bestimmte SNOMED-Indizes in einer „region“ bzw. in einer Regel, so muß untersucht werden, ob durch eine andere Regel ein anderer ICD-9-Kode zugeordnet wird.

Die mangelnde Konsistenz der ICD-9 ist nicht nur nachteilig für die manuelle Kodierung, sondern auch für algorithmische Lösungen. Es darf daher nicht wundern, daß bei der Transformation von SNOMED in ICD-9 Fehler auftreten, die analysiert und durch Änderungen der Regeln verbessert werden müssen. Das Konsistenz-Defizit der ICD-9 liegt nach den Erfahrungen von R. THURMAYR im Bereich der 3-stelligen Kodes. Dagegen ist bei der Erweiterung auf den 4-stelligen ICD-9-Kode mit keinen größeren Schwierigkeiten zu rechnen. Auch bei der in Abschnitt 2 angesprochenen Umstellung von der Bevorzugung des *-Kodes auf die Bevorzugung des +-Kodes in ICD-9 dürfte die durch SNOMED ermöglichte algorithmische Lösung der Kodierung sehr hilfreich sein.

Die hier beschriebenen Methoden zur Bestimmung der Tiefe und Konsistenz einer Klassifikation dürften dann voll wirksam werden, wenn eine Klassifikation für ein bestimmtes Anwendungsgebiet neu entwickelt wird, das Programm von WINGERT im Rahmen der Entwicklung eingesetzt wird und die Ergebnisse bei der Entwicklung der Klassifikation Berücksichtigung finden. Dies wird durch Einsatz des Programms bei der Verwaltungsberufsgenossenschaft Hamburg und der Entwicklung eines neuen BG-Schlüssels versucht.

5.2. Vergleich zweier Klassifikationen

Zu den Phrasen jeder Klassifikation können durch das Programm die zugehörigen SNOMED-Indizes erzeugt werden. Wird dies für die Phrasen zweier Klassifikationen durchgeführt und ordnet man dann die SNOMED-Indizes, dann kann z. B. ein Zunahmefaktor von einem 1. Schlüssel auf einen 2. Schlüssel in Bezug auf die Häufigkeit der Nennung von Organsystemen bestimmt werden. In Abb. 2 wurde die 2. Version des Operationsschlüssels nach SCHEIBE mit der 1. Version verglichen. Die Graphik zeigt eine starke Zunahme der topographischen Begriffe im Bereich von Organsystemen, die nicht zur Allgemeinchirurgie zählen. Insgesamt ist in der 2. Version die Anzahl unterschiedlicher topographischer Angaben etwa 4 mal häufiger als in der 1. Version. Daran erkennt man, daß der SCHEIBE-Schlüssel in der 2. Version von einem allgemein chirurgischen Operationsschlüssel auf einen Operationsschlüssel für alle chirurgisch tätigen Fächer erweitert wurde.

In der klinischen Dokumentation kommt es immer wieder vor, daß die benutzten Klassifikationen bzw. Schlüssel gewechselt oder zumindest geändert werden. Dies bedeutet häufig, daß auch die Daten von Patienten früherer Jahrgänge umkodiert werden sollen oder gar müssen. Ziel ist dann die Erstellung einer Umsetzungstabelle von einem Schlüssel 1 in den neuen Schlüssel 2. Eine solche Transformation, die nur auf Kenntnis der Schlüssel beruht, ist natürlich nur möglich, wenn der Schlüssel 2 eine Vergrößerung des Schlüssels 1 ist. Ist dies nicht der Fall, dann benötigt man für die Umkodierung zusätzlich die Klartexte. Die Weiterentwicklung eines Schlüssels ist meist eine Verfeinerung, mit Überschneidungen muß aber immer gerechnet werden. In einer Umsetzungstabelle sollen zu jedem Code des Schlüssels 1 alle Texte des alphabetischen Teils mit dem zugehörigen Code des Schlüssels 2 aufgeführt werden.

Eine solche Umsetzungstabelle setzt natürlich den gleichen Anwendungsbereich (z. B. Diagnose, Operation oder Histologie) voraus. Auch bei gleichem Anwendungsbereich erleichtert eine solche Umsetzungstabelle die Arbeit kaum, falls unterschiedliche medizinische Lehrmeinungen hinter beiden Schlüsseln stehen oder verschiedene Anwendungsziele in der Klassifikation konkurrieren. Anderenfalls lohnt sich die Arbeit.

Zunächst wird man durch Eingabe der in beiden Schlüsseln verwendeten Bezeichnungen je eine Datei mit SNOMED-Indizes und den zugehörigen Codes des jeweiligen Schlüssels

erzeugen. Beide Dateien werden dann nach den Codes des jeweiligen Schlüssels geordnet, und es werden dann die Überschneidungen der zugehörigen „SNOMED-regions“ analysiert.

Diese Analyse kann sehr leicht so formalisiert werden, daß sie mit Hilfe eines Programms durchgeführt wird. Das Ergebnis des Programms ist eine Umsetzungstabelle, die der Benutzer noch überarbeiten muß.

R. THURMAYR hat für die Umsetzung des Lokalisationscodes des Operationsschlüssels nach SCHEIBE (Allgemeiner Teil) von Version 1 auf Version 2 ein dafür geeignetes FORTRAN-Programm geschrieben und mit Erfolg eingesetzt.

6. Aktueller Stand und weitere Entwicklung

Die Erweiterungen von SNOMED, die Entwicklung der Algorithmen und die Programmierung waren von FRIEDRICH WINGERT weitestgehend allein durchgeführt worden. Obwohl SNOMED und Algorithmen veröffentlicht waren und eine Beschreibung des Programms vorlag, war die Weiterführung dieser Arbeiten nach seinem plötzlichen Tod äußerst schwierig.

Dies ist i. w. darin begründet, daß FRIEDRICH WINGERT die vielfältigen Aspekte und Zusammenhänge zwischen der Nomenklatur SNOMED und den Algorithmen zur Indexierung und Transformation in die ICD als einziger vollständig überblickt hat. Es gab und gibt sonst niemanden, der dazu in der Lage ist. Die Fortführung seiner Arbeiten wird daher auf Dauer nur in geändertem Rahmen möglich sein.

Die Arbeiten an der Nomenklatur und an dem Programm bedingen einander. SNOMED ist einerseits so umfangreich, daß sie ohne Unterstützung durch ein Programm wenig Anwendung finden dürfte. Andererseits führte die Anwendung des Programms dazu, daß neue Terme aufgenommen wurden und SNOMED dadurch eine gewisse Vollständigkeit erreicht hat. Änderungen im medizinischen Sprachgebrauch und die fachliche Fort- und Weiterentwicklung auf den verschiedenen Gebieten können durch Anwendungen des Programms sehr schnell erkannt und in der Nomenklatur berücksichtigt werden.

6.1. Arbeiten an SNOMED

FRIEDRICH WINGERT hat an SNOMED laufend gearbeitet, sie modifiziert und ergänzt. SNOMED, wie sie in [3] veröffentlicht ist, ist nicht mehr die Version, die in dem Programm verwendet wird. Diese (aktuelle) Version wird aber gegenwärtig in den USA und Canada überarbeitet, in die Sprachen Englisch und Französisch übertragen und weiter ergänzt und modifiziert (SNOMED III).

Die Nomenklatur bedarf an einigen Stellen dringender Ergänzungen, so muß etwa die Dimension „D“ durch Querverweise auf die anderen Dimensionen ergänzt werden. Seit dem Tod von FRIEDRICH WINGERT ruhen alle Arbeiten an der deutschsprachigen Nomenklatur SNOMED.

Um die klinische Relevanz der Nomenklatur SNOMED auf Dauer zu gewährleisten, darf die Standardisierung nicht durch (notwendig werdende) lokale Ergänzungen blockiert werden. Für jeden Sprachbereich (also auch für den deutschsprachigen Raum) muß eine zentrale Stelle die Pflege der Nomenklatur organisieren. Diese muß zudem Änderungen mit

den zentralen Stellen anderer Sprachbereiche abstimmen. Dies muß möglichst schnell in Angriff genommen werden, da eine baldige und enge Zusammenarbeit mit dem „College of American Pathologists“ anzustreben ist.

Jede Nomenklatur muß zudem dem aktuellen klinischen Sprachgebrauch entsprechen. Dazu müssen in Diskussionen mit und zwischen Fachvertretern der entsprechenden medizinischen Disziplinen die begrifflichen Inhalte von Diagnosen und Therapien präzisiert und u. U. korrigiert werden. Dies führt zu Vereinbarungen, welche Begriffe in der Nomenklatur neu aufzunehmen, welche ggf. zu streichen und welche zu modifizieren sind.

6.2. Weiterentwicklung der Programme

Die automatische „Standardisierung“ von medizinischen Texten ist die wichtigste Grundlage für die Entwicklung von Expertensystemen und Informationssystemen in der Medizin. Durch die Möglichkeiten der automatischen Indexierung wird die Konsistenz der Verschlüsselung, die die Grundvoraussetzung für die Auswertbarkeit und Vergleichbarkeit von medizinischen Texten ist, stark erhöht.

Das Programm zur automatischen Kodierung nach SNOMED und der Abbildung auf ICD wurden von WINGERT in PL/I programmiert. Die Programme stehen z. Z. nur für Großrechner zur Verfügung. H. GÖTTSCHE stellt mit Hilfe von N. OSADA zur Zeit eine C-Version her, die auf einem PC-AT laufen wird. Der Indexierungsteil (Generierung der SNOMED-Indizes) wird voraussichtlich im Sommer 1990 und der sich daran anschließende Transformationsteil in die (3-stellige) ICD-9 Ende 1990 fertiggestellt sein.

Auch bei diesen Arbeiten geht es primär darum, den von FRIEDRICH WINGERT hinterlassenen Entwicklungsstand des Programms aufzuarbeiten und zu analysieren. Es bleiben viele Aufgaben:

- Die Indexierungsalgorithmen müssen weiter verbessert werden. Es muß überprüft werden, ob durch Anwendung von Methoden der Syntax- und Kontextanalyse die Ergebnisse verbessert werden können. So werden auf dem jetzigen Stand lexikalische Informationen nur bei Groß- und Kleinschreibung im Deutschen verwendet.
- Die von WINGERT begonnene Erweiterung der Transformationstabellen auf den 4-stelligen ICD muß zu Ende geführt werden.
- Entsprechende Transformationstabellen für andere wichtige Klassifikationen (etwa die der Berufsgenossenschaften) sind zu entwickeln.

Anhang A: Beispiel für die Umsetzung von Diagnosen in SNOMED

D nicht transmuraler Hinterwandinfarkt

T33030	hinteres Myokard
M54870	nichttransmuraler Infarkt
I410	akuter Myokardinfarkt

D dislozierte Herzschrittmacherelektrode

M18566	Schrittmacherdislokation
E90412	Schrittmacherelektrode
I996	Komplikationen, die bestimmten näher bez. Maßnahmen eigentümlich sind

D akute Pankreatitis Alkohol bedingt

T59000	Pancreas
M41000	akute Entzündung
E55100	Alkohol
I577	Krkht. der Bauchspeicheldrüse

- D Netzhautablösung li.
 TXX610 Retina
 M31140 Ablation
 GX2021 sin.
 I361 Netzhautablösung u. -defekte
- D drittgradig offene Unterschenkel-Trümmerfraktur nach Autounfall
 T11730 Tibia
 T11740 Fibula
 M12290 offene Mehrfragmentfraktur
 FY3720 Opfer bei Verkehrsunfall
 GX1730 hochgradig
 I823 Fraktur der Tibia u. der Fibula
- D Lippen-Kiefer-Gaumen-Spalte
 T10170 Maxilla
 T51100 Palatum
 T52100 Labium sup.
 M21510 kongenit. Fissur
 I749 Gaumen- und Lippenspalte
- D Neugeborenes diabetischer Mutter
 F35700 Risikokind
 D24100 mütterlicher Diabetes
 I775 Endokrine u. metabolische Störung speziell des Fetus und Neugeborenen

Anhang B: Beispiel für die Umsetzung anderer medizinischer Texte in SNOMED

- O Lipase-Bestimmung im Serum
 TOX500 Blutserum
 F22130 Lipase
 P40200 quant. chem. Verfahren
- O Monitoring von EKG, Blutdruck, Blutgasen, Herzminutenvolumen
 P71690 Monitoring von EKG, Blutdruck, Blutgasen, HMV
- O Herzschrittmacherwechsel
 E90410 Herzschrittmacher
 P12790 Reimplantation
- O Antibiotikatherapie
 E72000 Antibiotika
 P00101 Therapie
- O ERCP (endoskopisch retrograde Cholangio-Pankreatikographie)
 T58000 extrahepatischer Gallengang
 T59010 Ductus pancreaticus
 P12830 retrograde Kontrastmittelinjektion mittels Katheter

O Argonlaser-Koagulation Netzhaut

TXX610	Retina
E93271	Argonlaser
P15090	Photokoagulation

O ACVB

T42000	Aorta
T43000	A. coronaria cordis
P14360	autogener Bypass

O Babcock Varizen-Stripping

T48000	Vena
M32600	Varixknoten
P15810	Stripping mit Ligatur

Literatur

1. CÔTÉ, R. A. (ed.) (1982): Systematized Nomenclature of Medicine (SNOMED). Ed2. Volume I: Numerical Index. Volume II: Alphabetic Index. Skokie III, College of American Pathologists
2. CÔTÉ, R. A., ROTHWELL D. J. (1989): The classification-nomenclature issues in medicine: a return to natural language. Med Inf. 14, 25 - 41
3. WINGERT, F. (1984): SNOMED. Systematisierte Nomenklatur der Medizin. Band 1: Numerischer Index. Band 2: Alphabetischer Index. Springer Berlin - Heidelberg - New York
4. WINGERT, F. (1985): Automated indexing of SNOMED. Methods Inf. Med. 24, 27 - 34
5. WINGERT, F. (1985): Morphologic Analysis of Compound Words. Methods Inf. Med. 24, 155 - 162
6. WINGERT, F. (1986): An Indexing System for SNOMED. Methods Inf. Med. 25, 22 - 30
7. WINGERT, F. (1987): Automated indexing of SNOMED statements into ICD. Methods Inf. Med. 26, 93 - 98.
8. Handbuch der Internationalen Klassifikation der Krankheiten, Verletzungen und Todesursachen (ICD), 9. Revision, Band I und II (1979). Consulting Verlag Wuppertal

Anschriften der Verfasser: Prof. Dr. J. Berger, Universität Hamburg, Institut für Mathematik und Datenverarbeitung in der Medizin, Martinistr. 52, D-2000 Hamburg 20 · PD Dr. E. Hultsch, Universität Münster, Institut für Medizinische Informatik und Biomathe-matik, Domagkstr. 9, D-4400 Münster · Prof. Dr. R. Thurmayr, TH München, Institut für Medizinische Statistik und Epidemiologie, Ismaninger Str. 22, D-8000 München 80